

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:49:28 ; Search time 43.8146 Seconds
(without alignments)
9625.515 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acagaactggcaaggca.....agggcctgtcaccacgga 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	51	100.0	365	3	US-09-925-299-668 Sequence 668, App
2	51	100.0	365	3	US-09-925-299-668 Sequence 668, App
3	51	100.0	1777	8	US-10-623-108-7 Sequence 7, Appli
4	51	100.0	1837	8	US-10-623-108-5 Sequence 5, Appli
5	50	98.0	13808	6	US-09-764-877-3351 Sequence 3351, Ap
6	50	98.0	13808	6	US-10-242-515-3351 Sequence 3351, Ap
7	25.6	50.2	9916	7	US-09-816-095-3 Sequence 3, Appli
8	25.6	50.2	9916	7	US-10-634-905-3 Sequence 275780,
9	24.8	48.6	506	5	US-10-027-632-275790 Sequence 275790,
10	24.8	48.6	506	6	US-10-027-632-275790 Sequence 650600,
11	24.8	48.6	594	4	US-09-925-065A-650600 Sequence 21955, A
12	24.8	48.6	992	8	US-10-357-930-21955 Sequence 27809, A
13	24.8	48.6	992	8	US-10-357-930-27809 Sequence 3069, Ap
14	24.4	47.8	331	3	US-09-764-877-3070 Sequence 3070, Ap
15	24.4	47.8	331	3	US-10-242-515-3069 Sequence 3070, Ap
16	24.4	47.8	331	6	US-10-242-515-3070 Sequence 1105, Ap
17	24.4	47.1	582	8	US-10-774-355A-1105 Sequence 97357, A
18	24	47.1	2112	5	US-10-027-632-97357 Sequence 97357, A
19	24	47.1	2112	5	US-10-027-632-97358 Sequence 97358, A
20	24	47.1	2112	5	US-10-027-632-97359 Sequence 97359, A
21	24	47.1	2112	5	US-10-027-632-103733 Sequence 103733,
22	24	47.1	2112	5	US-10-027-632-103733 Sequence 103733,
23	24	47.1	2112	5	US-10-027-632-103734 Sequence 103734,

C 24	24	47.1	2112	5	US-10-027-632-103735	Sequence 103735, A
C 25	24	47.1	2112	6	US-10-027-632-97357	Sequence 97357, A
C 26	24	47.1	2112	6	US-10-027-632-97358	Sequence 97358, A
C 27	24	47.1	2112	6	US-10-027-632-97359	Sequence 97359, A
C 28	24	47.1	2112	6	US-10-027-632-103733	Sequence 103733, A
C 29	24	47.1	2112	6	US-10-027-632-103734	Sequence 103734, A
C 30	24	47.1	2112	6	US-10-027-632-103735	Sequence 103735, A
C 31	24	47.1	51657	6	US-10-057-475B-10475	Sequence 10475, A
C 32	24	47.1	51657	6	US-10-154-884B-10475	Sequence 10475, A
C 33	23.8	46.7	346	8	US-10-425-115-173660	Sequence 173660, A
C 34	23.8	46.7	1312	5	US-10-106-698-2115	Sequence 2115, Ap
C 35	23.8	46.7	1562	5	US-10-027-632-255989	Sequence 255989, A
C 36	23.8	46.7	1562	6	US-10-027-632-255989	Sequence 255989, A
C 37	23.8	46.7	3863	5	US-10-103-313-132	Sequence 132, App
C 38	23.8	46.7	247682	7	US-10-235-192A-28	Sequence 28, Appli
C 39	23.6	46.3	303	6	US-10-006-285-5	Sequence 5, Appli
C 40	23.6	46.3	506	6	US-10-029-386-6668	Sequence 6668, Ap
C 41	23.6	46.3	7862	8	US-10-723-860-4981	Sequence 4981, Ap
C 42	23.6	46.3	64011	8	US-10-719-993-6991	Sequence 6991, Ap
C 43	23.6	46.3	85571	8	US-10-719-993-6991	Sequence 6991, Ap
C 44	23.4	45.9	405	8	US-10-425-115-33612	Sequence 33612, A
C 45	23.4	45.9	17177	6	US-10-225-810-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-925-299-668
Sequence 668, Application US/09925299
Request No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 668
LENGTH: 365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (172)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (239)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (243)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (244)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (329)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (330)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (353)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature

; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-299-668

Query Match 100.0%; Score 51; DB 3; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 32 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 82

RESULT 2

US-09-925-299-668
; Sequence 668, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 668
; LENGTH: 365

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (172)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (239)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (243)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (244)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (329)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (330)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (353)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-299-668

Query Match 100.0%; Score 51; DB 3; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 32 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 82

RESULT 3

US-10-623-108-7
; Sequence 7, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHWO
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-623-108-7

Query Match 100.0%; Score 51; DB 8; Length 1777;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 1126 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 1176

RESULT 4

US-10-623-108-5
; Sequence 5, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:

; APPLICANT: DAI, KEN-SHWO
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-623-108-5

Query Match 100.0%; Score 51; DB 8; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 1186 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 1236

RESULT 5

US-09-764-877-3351
; Sequence 3351, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3351

```

; LENGTH: 13808
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3351

Query Match      98.0%; Score 50; DB 3; Length 13808;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 50
    |||||
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 6424

RESULT 6
US-10-242-515-3351
; Sequence 3351, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3351
; LENGTH: 13808
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-3351

Query Match      98.0%; Score 50; DB 6; Length 13808;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 50
    |||||
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 6424

RESULT 7
US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weinu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095

```

```

; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match      50.2%; Score 25.6; DB 3; Length 99916;
Best Local Similarity 70.8%; Pred. No. 6.2;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAG 49
    |||||
Db 1819 CAGACCCGGGAGAGCCAAAGGGTCCCTCCGGGCTCCCTCTCGCAG 1866

RESULT 8
US-10-634-905-3
; Sequence 3, Application US/10634905
; Publication No. US20040067225A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weinu et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001147-DIV
; CURRENT APPLICATION NUMBER: US/10/634,905
; CURRENT FILING DATE: 2003-08-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-10-634-905-3

Query Match      50.2%; Score 25.6; DB 7; Length 99916;
Best Local Similarity 70.8%; Pred. No. 6.2;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAG 49
    |||||
Db 1819 CAGACCCGGGAGAGCCAAAGGGTCCCTCCGGGCTCCCTCTCGCAG 1866

RESULT 9
US-10-027-632-275790
; Sequence 275790, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275790
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-275790

Query Match      48.6%; Score 24.8; DB 5; Length 506;
Best Local Similarity 72.7%; Pred. No. 8;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      8  TGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      105 TGGCAAGGTGCCCAAGAGTCCTTAATCGCATCTGTCTCCAGGA 148

RESULT 10
US-10-027-632-275790
; Sequence 275790, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275790
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-275790

Query Match      48.6%; Score 24.8; DB 6; Length 506;
Best Local Similarity 72.7%; Pred. No. 8;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      8  TGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      105 TGGCAAGGTGCCCAAGAGTCCTTAATCGCATCTGTCTCCAGGA 148

RESULT 11
US-09-925-065A-650600
; Sequence 650600, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
```

```
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 650600
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-650600

Query Match      48.6%; Score 24.8; DB 4; Length 594;
Best Local Similarity 72.7%; Pred. No. 8.1;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      8  TGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      85  TGGCAAGGTGCCCAAGAGTCCTTAATCGCATCTGTCTCCAGGA 128

RESULT 12
US-10-357-930-21955
; Sequence 21955, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21955
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: 992
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-21955

Query Match      48.6%; Score 24.8; DB 8; Length 992;
Best Local Similarity 72.7%; Pred. No. 8.5;
```


This Page Blank (uspto)

ORIGIN

```
Query Match      100.0%; Score 51; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.6e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 51
    |||||||
Db 309 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 359

RESULT 2
LOCUS BM702058 555 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-CQ1-aeY-e-05-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone
            UI-E-CQ1-aeY-e-05-0-UI 5', mRNA sequence.
ACCESSION  BM702058
VERSION     BM702058.1 GI:19015316
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE   1 (bases 1 to 555)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
PUBMED     889548
COMMENT    Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source           1..555
     mol_type="mRNA"
     db_xref="taxon:9606"
     clone="UI-E-CQ1-aeY-e-05-0-UI"
     tissue_type="optic nerve"
     dev_stage="adult"
     lab_host="DHI0B (Life Technologies) (T1 phage resistant)"
     clone_lib="UI-E-CQ1"
     notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
            modified polylinker; Site 1: EcoR I; Site 2: Not I;
            UI-E-CQ1 is a normalized cDNA library containing the
            following tissue(s): optic nerve. The library was
            constructed according to Bonaldo, Lennon and Soares,
            Genome Research, 6:791-806, 1996. First strand cDNA
            synthesis was primed with an oligo-dT primer containing a
            Not I site. Double stranded cDNA was ligated to an EcoR I
            adaptor, digested with Not I, and cloned directionally
            into pT7T3-Pac vector. The oligonucleotide used to prime
            the synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            (dT)18 tail. The sequence tag for this library is
            CCATTAAGTG. This library was created for the program, Gene
            Discovery in the Visual System, supported by National Eye
            Institute (NEI)."
```

ORIGIN

```
Query Match      100.0%; Score 51; DB 3; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
```

```
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 51
    |||||||
Db 356 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 406

RESULT 3
LOCUS CB156158 575 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0214813 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-2-C09
            5', mRNA sequence.
ACCESSION  CB156158
VERSION     CB156158.1 GI:28141276
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE   1 (bases 1 to 575)
AUTHORS    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 2 row: C column: 09
            High quality sequence stop: 575.
FEATURES             Location/Qualifiers
     source           1..575
     mol_type="mRNA"
     db_xref="taxon:9606"
     clone="L18POOL1n1-2-C09"
     cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
     lab_host="T0p10P"
     clone_lib="L18POOL1n1"
     note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
            Site 2: NotI; The library was contributed by The Soares
            laboratory and it was constructed as described by Bonaldo,
            M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
            6(9): 791-806. RNA was prepared from harvested cell
            culture."
```

ORIGIN

```
Query Match      100.0%; Score 51; DB 6; Length 575;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 51
    |||||||
Db 222 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 272
```

RESULT 4

```
BE891264
LOCUS BE891264 653 bp mRNA linear EST 20-OCT-2000
DEFINITION 601432092F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3917255 5',
            mRNA sequence.
ACCESSION  BE891264
VERSION     BE891264.1 GI:10350423
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```

REFERENCE
AUTHORS      1 (bases 1 to 653)
TITLE        NIH-MGC http://mgi.nci.nih.gov/
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabs-r@mail.nih.gov
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1M9743 row: h column: 24
              High quality sequence stop: 602.

FEATURES
source
1..653
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:3917255"
   /tissue_type="melanotic melanoma"
   /lab_host="DH10B (phage-resistant)"
   /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
   Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
   Average insert size 2 kb. Library constructed by Life
   Technologies."

ORIGIN
Query Match      100.0%; Score 51; DB 2; Length 653;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
Db 416 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 466

RESULT 5
BM041252
LOCUS      6031614446F1 NIH_MGC_108 Homo sapiens cDNA clone IMAGE:5420314 5',
DEFINITION mRNA sequence.
ACCESSION  BM041252
VERSION     BM041252.1 GI:16770519
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.

REFERENCE
1 (bases 1 to 673)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1874 row: 1 column: 11
High quality sequence stop: 648.

FEATURES
source
1..673
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 51; DB 2; Length 653;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
Db 416 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 466

RESULT 5
BM041252
LOCUS      6031614446F1 NIH_MGC_108 Homo sapiens cDNA clone IMAGE:5420314 5',
DEFINITION mRNA sequence.
ACCESSION  BM041252
VERSION     BM041252.1 GI:16770519
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.

REFERENCE
1 (bases 1 to 673)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1874 row: 1 column: 11
High quality sequence stop: 648.

FEATURES
source
1..673
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"

```

```

/clone="IMAGE:5420314"
/tissue_type="Wilms' tumor, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 108"
/note="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 51; DB 3; Length 673;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
Db 161 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 211

RESULT 6
CA446812/c
LOCUS      673 bp mRNA linear EST 08-NOV-2002
DEFINITION UI-H-EDI-axx-p-03-0-UI.s1 NCI CGAP ED1 Homo sapiens cDNA clone
            UI-H-EDI-axx-p-03-0-UI 3', mRNA sequence.
ACCESSION  CA446812
VERSION     CA446812.1 GI:24811232
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.

REFERENCE
1 (bases 1 to 673)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1..673
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="UI-H-EDI-axx-p-03-0-UI"
   /tissue_type="Chondrosarcoma"
   /dev_stage="Adult"
   /lab_host="DH10B (Life Technologies)"
   /clone_lib="NCI CGAP ED1"
   /note="Organ: Left Pubic Bone; Vector: pT7T3-Pac
   (Pharmacia) with a modified polylinker; Site 1: EcoR I;
   Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library
   containing the following tissue(s): Chondrosarcoma cell
   line C55. The library was constructed according to
   Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
   1996. First strand cDNA synthesis was primed with an
   oligo-dT primer containing a Not I site. Double stranded
   cDNA was ligated to an EcoR I adaptor, digested with Not
   I, and cloned directionally into pT7T3-Pac vector. The
   oligonucleotide used to prime the synthesis of
   first-strand cDNA contains a library tag sequence that is

```

located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-ED1
TAG_SEQ=CGTCAAGGCT"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 673;
Best Local Similarity 100.0%; Pred. No. 3.7e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
|||||
670 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 620

Db

RESULT 7

BM041386/c
LOCUS 773 bp mRNA linear EST 07-NOV-2001
DEFINITION 603614446T1 NIH_MGC_108 Homo sapiens cDNA clone IMAGE:5420314 3',
mRNA sequence.

ACCESSION BM041386

VERSION BM041386.1 GI:16770653

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 773)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM1874 row: 1 column: 11

High quality sequence start: 26

High quality sequence stop: 718.

FEATURES

source

1..773
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5420314"
/tissue_type="Wilms' tumor, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 108"
/notes="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 51; DB 3; Length 773;
Best Local Similarity 100.0%; Pred. No. 3.8e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
|||||
611 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 561

Db

RESULT 8

AL576621/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Al576621.3 GI:46255607

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 864)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 16, 2001 this sequence version replaced gi:31314903.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

8342.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 51; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 3.8e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
|||||
616 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 566

Db

RESULT 9

B50582/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

B50582.1 GI:2602819

GSS.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 307)

AUTHORS Kim U.-J., Adams, M.D. and Simon, M.I.

TITLE Determination of clone end sequences of human Bacterial Artificial

Chromosomes

JOURNAL Unpublished (1997)

COMMENT Contact: Ung-Jin Kim

Caltech Genome Research Lab

AL576621 864 bp mRNA linear EST 07-APR-2004
clone CSODI078YM11 3-PRIME, mRNA sequence.

AL576621.3 GI:46255607

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 864)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 16, 2001 this sequence version replaced gi:31314903.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

8342.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

```
/clone_lib="Human Retina cDNA (Un-normalized,
```

ORIGIN

Query Match 73.3%; Score 37.4; DB 2; Length 855;
 Best Local Similarity 96.1%; Pred. No. 0.024; 1; Indels 1; Gaps 1;
 Matches 49; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
 |||||
 Db 572 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGA-GGCCTCTGTCCACCAGGA 621
 |||||

RESULT 12

AA324242
 LOCUS
 DEFINITION EST27116 Cerebellum II Homo sapiens cDNA 5' end similar to EST
 containing Alu repeat, mRNA sequence.

ACCESSION AA324242

VERSION AA324242.1 GI:1976559

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.

REFERENCE 1 (bases 1 to 500)

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fritchugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
 Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Falange,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Peng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissenner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

7566098

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

FEATURES

1..500
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):124792"
 /db_xref="taxon:9606"
 /tissue_type="cerebellum"
 /dev_stage="adult"
 /clone_lib="Cerebellum II"
 /notes="Organ: brain; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

ORIGIN

Query Match 55.7%; Score 28.4; DB 1; Length 500;
 Best Local Similarity 80.0%; Pred. No. 34;
 Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY

8 TGGCAAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACC 47
 |||||

Db

366 TGGCAAAGAGGCAAGAGGTCATTTGAGGGGCTTTTNNACC 405
 |||||

RESULT 13

CV358408

LOCUS

DEFINITION PMO-AN0087-270101-008-g09 AN0087 Homo sapiens cDNA, mRNA sequence.

ACCESSION CV358408

VERSION CV358408.1 GI:52708463

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.

REFERENCE 1 (bases 1 to 412)

AUTHORS

Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. <http://www.ludwig.org.br>.

FEATURES

source

1..412
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="AN0087"
 /note="Organ: amnion normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196.716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 54.5%; Score 27.8; DB 7; Length 412;
 Best Local Similarity 74.5%; Pred. No. 53;
 Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY

4 GAACGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGG 50
 |||||

Db

46 GAACGGCAAGAGGCTAAAGGTCACCGTGTGTCTGTGGCTCCAGG 92
 |||||

RESULT 14

BE385462

LOCUS

DEFINITION 601276049F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3616813 5',

EST.

ACCESSION BE385462

VERSION BE385462.1 GI:9330827

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 680)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLM283 row: b column: 14
High quality sequence stop: 617.

FEATURES
source
1. .680
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3616813"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 20"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match 52.9%; Score 27; DB 2; Length 680;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 ACAGACTGGCAAGAGGCAAGAGGTCTACTGAGGGCTCTGTCTACCCAGGA 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 ACAGAGGTCTCAAGACCCAGAGACTCTGGGAGGCACTGCCACCCAGGA 317

RESULT 15
BI869166/c
LOCUS
DEFINITION BI869166 Homo sapiens 698 bp mRNA linear EST 11-OCT-2001
603395496F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405446 5',
mRNA sequence.
ACCESSION BI869166
VERSION BI869166.1 GI:16042826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 698)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM12033 row: p column: 23

FEATURES
source
High quality sequence stop: 539.
Location/Qualifiers
1. .698
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405446"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 52.2%; Score 26.6; DB 3; Length 698;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 3 AGAACTGGCAAGAGGCAAGAGGTCTACTGAGGGCTCTGTCTACCCAGGA 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 650 AGAAGGGGTAAATAGTGAAGGGGAGCTGAGGCTTGAGTCACCCAGGA 602

Search completed: February 7, 2006, 14:53:04
Job time : 204.677 secs

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:22:44 ; Search time 31.0376 Seconds
(without alignments)
10951.209 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acagaactggcaaggca.....agggctctgtcaccaggga 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- N_Geneseq_21:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001as:*
 - 5: Geneseq2001bs:*
 - 6: Geneseq2002as:*
 - 7: Geneseq2002bs:*
 - 8: Geneseq2003as:*
 - 9: Geneseq2003bs:*
 - 10: Geneseq2003cs:*
 - 11: Geneseq2003ds:*
 - 12: Geneseq2004as:*
 - 13: Geneseq2004bs:*
 - 14: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	365	3 AAC98658	AAC98658 Human col
2	51	100.0	1210	13 ACN41203	ACN41203 Human dia
3	51	100.0	1290	13 ACN41200	ACN41200 Human dia
4	51	100.0	1305	13 ACN41199	ACN41199 Human dia
5	51	100.0	1777	14 ADW12906	ADW12906 Human SMA
6	51	100.0	1837	14 ADW12904	ADW12904 Human SMA
7	50	98.0	13808	8 AAL36986	AAL36986 Human mus
8	50	98.0	13808	8 ABX59974	ABX59974 cDNA enco
9	50	98.0	13808	12 ADJ30724	ADJ30724 Human mus
10	25.6	50.2	99916	6 ADI03931	ADI03931 Human enz
11	25.6	50.2	116297	12 ADQ97587	ADQ97587 Human can
12	24.8	48.6	992	5 ABV21964	ABV21964 Human pro
13	24.8	48.6	992	5 ABV27796	ABV27796 Human pro
14	24.4	47.8	331	4 AAL36705	AAL36705 Human mus
15	24.4	47.8	331	4 AAL36704	AAL36704 Human mus
16	24.4	47.8	331	8 ABX59692	ABX59692 cDNA enco
17	24.4	47.8	331	8 ABX59693	ABX59693 cDNA enco
18	24.4	47.8	331	12 ADJ30443	ADJ30443 Human mus
19	24.4	47.8	331	12 ADJ30442	ADJ30442 Human mus

C	20	24.4	47.8	1023	14	ACL73128	ACL73128 M. xanthu
C	21	24.4	47.8	13346	14	ACL64626	ACL64626 M. xanthu
C	22	23.8	46.7	1008	3	AAC74422	AAC74422 Human sec
C	23	23.8	46.7	1008	8	ABZ73645	ABZ73645 Secreted
C	24	23.8	46.7	1008	8	ADA98137	ADA98137 Human sec
C	25	23.8	46.7	1008	10	ABZ67239	ABZ67239 Human sec
C	26	23.8	46.7	1043	3	AAC93331	AAC93331 Human sec
C	27	23.8	46.7	1312	4	AAS35023	AAS35023 Human col
C	28	23.8	46.7	2140	13	ADU82622	ADU82622 Human MDD
C	29	23.8	46.7	3863	4	AAS34888	AAS34888 cDNA enco
C	30	23.8	46.7	3863	10	ADC46046	ADC46046 Human neo
C	31	23.8	46.7	247682	12	ADL08109	ADL08109 Human gen
C	32	23.6	46.3	303	10	ADI22195	ADI22195 Rat liver
C	33	23.6	46.3	7862	12	ACH73473	ACH73473 Human gen
C	34	23.6	46.3	506	12	ADQ22161	ADQ22161 Human sof
C	35	23.4	45.9	2932	8	ADA41546	ADA41546 Human sec
C	36	23.4	45.9	2932	10	ADA57677	ADA57677 BAC fragm
C	37	23.4	45.9	17173	8	ABZ80234	ABZ80234 Mouse tra
C	38	23.2	45.5	1756	11	ADM01710	ADM01710 Human CDN
C	39	23.2	45.5	3089	10	ADB62760	ADB62760 Human CDN
C	40	23.2	45.5	3356	13	ADR08430	ADR08430 Full leng
C	41	23.2	45.5	11706	5	ABA20479	ABA20479 Human ner
C	42	23.2	45.5	28564	10	ADD47140	ADD47140 Human gen
C	43	23.2	45.5	34875	10	ADR14677	ADR14677 Human src
C	44	23.2	45.5	34875	13	ADR52965	ADR52965 Drug ther
C	45	23	45.1	556	4	AAL18474	AAL18474 Probe #84

ALIGNMENTS

RESULT 1
AAC98658
ID AAC98658 standard; cDNA; 365 BP.

AC AAC98658;
XX 09-MAR-2001 (first entry)
XX Human colon cancer antigen nucleotide sequence SEQ ID NO:668.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX identification; cytostatic; cardioactive; neuroprotective; vulnerar;
XX immunomodulatory; muscular; gynaecological; gastrointestinal;
XX nephrotropic; anti-infective; antibacterial; gene therapy; wound;
XX neural disorder; immune system disorder; muscular disorder;
XX reproductive disorder; gastrointestinal disorder; renal disorder;
XX infectious disease; cardiovascular disorder; 88.

OS Homo sapiens.
XX WO200055351-A1.
XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005883.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587534/55.

XX P-PSDB; AAB53901.

XX Colon cancer associated gene sequences, referred to as colon cancer
XX antigens, useful for the treatment, prevention, and diagnosis of colon
XX disorders such as colon cancer.

XX Claim 1; Page 1200; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,

called human colon cancer antigens, given in AA53234 to AA54006. The human colon cancer antigens can have cytostatic, cardioactive, muscular, neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and vulvular, nephrotropic, antineoplastic and antibacterial activities, and can be used in gene therapy. The colon cancer antigen polynucleotides, proteins and antibodies to the proteins are useful for the prevention, treatment and diagnosis of colon disorders, such as colon cancer. The polynucleotides may be used in diagnostics and research, such as for chromosome identification, and as hybridisation probes. The proteins may also be used to prevent diseases such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, wounds, renal disorders, infectious diseases, and cardiovascular disorders. AA98764 to AAC98772 and AA54007 represent sequences used in the exemplification of the present invention

Sequence 365 BP: 54 A; 121 C; 101 G; 79 T; 0 U; 10 Other;

Query Match	Score 51	DB 3	Length 365
Best Local Similarity	100.0%	Pred. No. 2.9e-09	
Matches 51	Conservative	0	Mismatches 0
		Indels 0	Gaps 0

QY 1 ACAGAACTGGCAAGAAGCGCAAGAGGTCACTGAGGGCCTCTGTCACCCAGGA 51
|||||
dh 32 ACAGAACTGGCAAGAAGCGCAAGAGGTCACTGAGGGCCTCTGTCACCCAGGA 82
|||||

RESULT 2
ACN41203
ID ACN41203 standard; cDNA; 1210 BP.
XX
XX AC ACN41203;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:78.
XX
XX KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX
XX KW dithp.
XX
XX OS Homo sapiens.
XX
XX PN WO2004023973-A2.
XX
XX PD 25-MAR-2004.
XX
XX PF 12-SEP-2003; 2003WO-US028227.
XX
XX PR 12-SEP-2002; 2002US-0410259P.
XX
XX PR 12-SEP-2002; 2002US-0410260P.
XX
XX PA (INCY-) INCYTE CORP.

xx New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.

PS Claim 1: Page: 190pp; English.

CC The invention relates to novel diagnostic and therapeutic polynucleotides
AA
CC selected from one of the 2722 sequences defined in the specification. A
CC

CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dthp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dthp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIGO at www.wigo.int/pct/en/sequences/listing.htm

Sequence 1210 BP: 248 A; 375 C; 376 G; 211 T; 0 U; 0 Other;

Query Match	100.0%;	Score 51;	DB 13;	Length 1210;
Best Local Similarity	100.0%;	Pred. No. 3.8e-09;		
Matches 51;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps 0;

Qy 1 ACAGAACTGGCAAAGAGGCAAGAGGTCACTGAGGGCCTCTGTACCCAGGA 51
1024 ACAGAACTGGCAAAGAGGCAAGAGGTCACTGAGGGCCTCTGTACCCAGGA 1074
Db

RESULT_3	
ACN41200	ACN41200 standard; cDNA; 1290 BP.
ID	ACN41200 standard; cDNA; 1290 BP.
XX	XX
XX	AC ACN41200;
XX	XX
DT	18-NOV-2004 (first entry)
XX	XX
DE	Human diagnostic and therapeutic polynucleotide SEQ ID NO:75.
XX	ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW	dithp.
KW	XX
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO2004023973-A2.
XX	XX
PD	25-MAR-2004.
XX	XX
PF	12-SEP-2003; 2003WO-US028227.
XX	XX
PR	12-SEP-2002; 2002US-0410259P.
PR	12-SEP-2002; 2002US-0410260P.
XX	XX
PA	(INCY-) INCYTE CORP.

New diagnostic and therapeutic polynucleotides and polypeptides, useful
 in diagnosing a condition, disease or disorder associated with human
 molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 in gene mapping.

PS Claim 1; Page; 190pp; English.

The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A

CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorders, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp polynucleotide of
 CC the invention. Note: The sequence data for this patent is not represented
 CC in the printed specification, but was obtained in electronic format
 CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 1290 BP; 261 A; 402 C; 398 G; 229 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 13; Length 1290;
 Best Local Similarity 100.0%; Pred. NO. 3.8e-09;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCCAGGA 51
 |||||
 DB 1104 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCCAGGA 1154

RESULT 4
 ACN41199
 ID ACN41199 standard; cDNA; 1305 BP.
 XX
 AC ACN41199;
 DT 18-NOV-2004 (first entry)
 XX
 DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:74.

ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
 dithp.

OS Homo sapiens.
 XX WO2004023973-A2.
 XX
 XX 25-MAR-2004.
 XX
 XX 12-SEP-2003; 2003WO-US028227.
 XX
 XX 12-SEP-2002; 2002US-0410259P.
 XX
 XX 12-SEP-2002; 2002US-0410260P.
 XX
 XX (INCY-) INCYTE CORP.
 XX
 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 XX Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
 XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 XX Peralta CH, Anderson SB, Rioux P, Shen RJ, Wu MC, Stuve LH;
 XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
 XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 XX Patury S, Shi X, Suarez CJ;

WPI; 2004-329368/30.
 DR P-PSDB; ABM82547.
 XX

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.

XX Claim 1; Page: 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A

CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorders, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp polynucleotide of
 CC the invention. Note: The sequence data for this patent is not represented
 CC in the printed specification, but was obtained in electronic format
 CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX

SQ Sequence 1305 BP; 273 A; 411 C; 392 G; 229 T; 0 U; 0 Other;
 Query Match 100.0%; Score 51; DB 13; Length 1305;
 Best Local Similarity 100.0%; Pred. NO. 3.8e-09;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCCAGGA 51
 |||||
 DB 1119 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCCAGGA 1169

RESULT 5
 ADW12906
 ID ADW12906 standard; cDNA; 1777 BP.
 XX
 AC ADW12906;
 XX
 DT 07-APR-2005 (first entry)
 XX
 DE Human SNAPK3V4 variant gene.
 XX
 DE DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
 XX gene; ss.

OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 12..1091
 XX /*tag= a
 XX /product= "SNAPK3V4 protein"

XX US2005013817-A1.
 XX
 XX 20-JAN-2005.
 XX
 XX 18-JUL-2003; 2003US-00623108.
 XX
 XX 18-JUL-2003; 2003US-00623108.
 XX
 XX (DAIK/) DAI K.
 XX
 XX Dai K;
 XX WPI; 2005-080923/09.
 XX P-PSDB; ADW12907.

XX New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
 PT diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
 PT in a mammal.

XX Claim 4; SEQ ID NO 7; 55pp; English.

XX The invention relates to novel isolated polypeptide (I) comprising an
 CC amino acid sequence selected from sequences comprising 335 or 359 amino
 CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
 CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
 CC associated with the deficiency of the SNAPK3 gene in a mammal. The
 CC fragments of the polypeptides and polynucleotides can also be used as

CC primers or probes. This sequence corresponds to the cDNA encoding the
 CC SNAPK3V1 variant protein.

SQ Sequence 1777 BP; 355 A; 575 C; 507 G; 340 T; 0 U; 0 Other;
 Query Match 100.0%; Score 51; DB 14; Length 1777;
 Best Local Similarity 100.0%; Pred. No. 4.1e-09;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
 |||||
 Db 1126 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 1176

RESULT 6
 ADW12904
 ID ADW12904 standard; cDNA; 1837 BP.

XX AC ADW12904;
 XX
 DT 07-APR-2005 (first entry)
 XX
 DE Human SNAPK3V3 variant gene.
 XX
 KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 12..1151
 FT /*tag= a
 FT /product= "SNAPK3V3 protein"

XX US2005013817-A1.
 XX 20-JAN-2005.
 XX 18-JUL-2003; 2003US-00623108.
 XX 18-JUL-2003; 2003US-00623108.
 XX (DAIK/) DAI K.
 XX Dai K;
 XX WPI; 2005-080923/09.
 XX P-PSDB; ADW12905.
 XX New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
 PT diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
 PT in a mammal.

XX Claim 4; SEQ ID NO 5; 55pp; English.
 XX
 CC The invention relates to novel isolated polypeptide (I) comprising an
 CC amino acid sequence selected from sequences comprising 335 or 359 amino
 CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
 CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
 CC associated with the deficiency of the SNAPK3 gene in a mammal. The
 CC fragments of the polypeptides and polynucleotides can also be used as
 CC primers or probes. This sequence corresponds to the cDNA encoding the
 CC SNAPK3V1 variant protein.

SQ Sequence 1837 BP; 369 A; 593 C; 526 G; 349 T; 0 U; 0 Other;
 Query Match 100.0%; Score 51; DB 14; Length 1837;
 Best Local Similarity 100.0%; Pred. No. 4.1e-09;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
 |||||
 Db 1186 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 1236

RESULT 7

AA136986
 ID AAL36986 standard; DNA; 13808 BP.

XX AAL36986;
 AC
 XX
 DT 08-JAN-2002 (first entry)
 XX
 DE Human musculoskeletal system related polynucleotide SEQ ID NO 3351.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200155367-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001338.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 18-AUG-2000; 2000US-0225759P.
 PR 22-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.

08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-023497P.
25-SEP-2000; 2000US-023498P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236357P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246612P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.

17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
05-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX Example 2; SEQ ID NO 3351; 781pp + Sequence Listing; English.
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (AB03087-AB04109) associated with the musculoskeletal system useful for
XX preventing, treating or ameliorating medical conditions e.g. By protein
XX or gene therapy. The genes are isolated from a range of human tissues
XX disclosed in the specification. The nucleic acids, proteins, antibodies
XX and (ant)agonists are useful in the diagnosis, treatment and prevention
XX of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung, or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 13808 BP; 2990 A; 3953 C; 3656 G; 3209 T; 0 U; 0 Other;
Query Match 98.0%; Score 50; DB 4; Length 13808;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCCTCTGTACCCAGG 50
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCCTCTGTACCCAGG 6424
RESULT 8
ABX59974
ID ABX59974 standard; cDNA; 13808 BP.
XX AC ABX59974;
XX 26-FEB-2003 (first entry)
XX cDNA encoding novel human musculoskeletal system antigen #2318.
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX

post-operative tissue repair; limb regeneration; neuronal growth;
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
AIDS-related complex; chondrocyte growth; bone regeneration;
periodontal regeneration; tissue transport; bone graft; skin aging;
keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
cell growth; organ transplant; cell differentiation; body height; weight;
hair colour; eye colour; skin; percentage of adipose tissue;
pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
depression; tendency for violence; pain; reproductive capability;
hormone level; endocrine level; appetite; libido; memory; stress;
storage capability; fat content; lipid content; protein content;
carbohydrate content; vitamin content; cofactor content;
nutritional component.

Homo sapiens.

US2002147140-A1.

10-OCT-2002.

17-JAN-2001; 2001US-00764877.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

28-JUN-2000; 2000US-0214886P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-0216880P.

11-JUL-2000; 2000US-0217487P.

11-JUL-2000; 2000US-0217496P.

14-JUL-2000; 2000US-0218290P.

26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225707P.

14-AUG-2000; 2000US-0225447P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.

22-AUG-2000; 2000US-0226868P.

30-AUG-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229287P.

01-SEP-2000; 2000US-0229343P.

01-SEP-2000; 2000US-0229344P.

01-SEP-2000; 2000US-0229345P.

05-SEP-2000; 2000US-0229509P.

05-SEP-2000; 2000US-0229513P.

08-SEP-2000; 2000US-0231413P.

21-SEP-2000; 2000US-0234223P.

21-SEP-2000; 2000US-0234223P.
25-SEP-2000; 2000US-0234997P.
27-SEP-2000; 2000US-0235834P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
17-NOV-2000; 2000US-0249299P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.

PA

PA

PA

XX

PI

XX

DR

XX

XX

PT

PT

XX

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-128199/12.

Isolated nucleic acid molecules encoding musculoskeletal system

associated polypeptides, useful for detecting disorders, e.g. cancer.

Disclosure; SEQ ID NO 3351; 321pp; English.

The invention describes an isolated nucleic acid molecule comprising a

sequence encoding musculoskeletal system associated polypeptides useful

for detecting disorders, e.g., cancer or cancer metastases, in animals or

humans. The nucleic acid stimulates re-vascularisation of ischaemic

tissues associated with conditions such as thrombosis, arteriosclerosis,

and other cardiovascular conditions; treats wounds due to injuries,

burns, post-operative tissue repair, and ulcers; stimulates angiogenesis

and limb regeneration; stimulates neuronal growth; can treat and prevent

neuronal damage occurring in certain disorders or neurodegenerative

conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-

related complex; stimulates chondrocyte growth, thus they can be used to

enhance bone and periodontal regeneration and aid in tissue transports or

bone grafts; prevents skin aging due to sunburn by stimulating

keratinocyte growth; prevents hair loss, since EGF family members

activate hair-forming cells and promotes melanocyte growth; stimulates

growth and differentiation of hematopoietic cells and bone marrow cells

when used in combination with other cytokines; maintains organs before

transplantation or for supporting cell culture of primary tissues;

induces tissue of mesodermal origin to differentiate in early embryos;

increases or decreases the differentiation or proliferation of embryonic

stem cells, besides, haematopoietic lineage; modulates mammalian

characteristics, such as, body height, weight, hair colour, eye colour,

skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,

cosmetic surgery); modulates mammalian metabolism; changes mammal's metal

state or physical state by influencing biorhythms, circadian rhythms,

depression, tendency for violence, tolerance for pain, reproductive

capabilities, hormonal or endocrine levels, appetite, libido, memory, or

stress; increases or decreases storage capabilities, fat content, lipid,

protein, carbohydrate, vitamins, minerals, cofactors or other nutritional

components. This sequence encodes a novel human musculoskeletal system

antigen. Note: The sequence data for this patent did not form part of the

printed specification, but was obtained in electronic format directly

from the US patent office at

ftp.segdata.uspto.gov/sequence.html?DocID=20020147140

Sequence 13808 BP; 2990 A; 3953 C; 3656 G; 3209 T; 0 U; 0 Other;

Query Match 98.0%; Score 50; DB 8; Length 13808;

Best Local Similarity 100.0%; Pred. No. 1.5e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGCGCTCTGTCAACCAGG 50

DB 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGCGCTCTGTCAACCAGG 6424

RESULT 9

ADJ30724

ID ADJ30724 standard; DNA; 13808 BP.

AC ADJ30724;

XX 20-MAY-2004 (first entry)

XX Human musculoskeletal system-associated genomic DNA - SEQ ID 3351.

DE musculoskeletal system; cytostatic; osteopathic; cancer; osteoporosis;

XX gene therapy; vaccine; human; ds.

XX Homo sapiens.

XX US2004009488-A1.
XX 15-JAN-2004.
XX 13-SEP-2002; 2002US-00242515.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764877.
PR (HUMA-) HUMAN GENOME SCI INC.
XX PA

XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2004-090458/09.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g., cancer of
PT musculoskeletal tissues or osteoporosis.
XX Disclosure; SEQ ID NO 3351; 289pp; English.
XX The invention relates to a novel isolated musculoskeletal system-
CC associated nucleic acid molecule. The nucleic acid of the invention
CC demonstrates cytostatic and osteopathic activities and may be useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition such as cancer of the musculoskeletal tissues or osteoporosis,
CC possibly via gene therapy or vaccine production. The current sequence is
CC that of the human musculoskeletal system-associated genomic DNA of the
CC invention. The current sequence is not shown within the specification per
CC se but is available on the USPTO web-site
CC http://seqdata.uspto.gov/sequence.html?DocID=20040009488.
XX
SQ Sequence 13808 BP; 2990 A; 3953 C; 3656 G; 3209 T; 0 U; 0 Other;
Query Match 98.0%; Score 50; DB 12; Length 13808;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 50
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 6424
RESULT 10
ADI03931
ID ADI03931 standard; DNA; 99916 BP.
XX
AC ADI03931;
XX
DT 22-APR-2004 (first entry)
DE Human enzyme protein encoding genomic DNA.
XX
DE Enzyme protein; drug screening; disease diagnosis; human; gene therapy;
KW chromosome 6; gene; ds; glucuronyltransferase.
XX
XX Homo sapiens.
XX Location/Qualifiers
FH 2614..96915
FT CDS
FT /*tag= b
FT /product= "enzyme protein"
FT /note= "contains introns"
FT 2614..3204
FT exon
FT /*tag= a
FT /number= 1
FT 3205..64312
FT intron
FT /*tag= c
FT /number= 1
FT 64313..64457
FT exon
FT /*tag= d
FT /number= 2
FT 64458..96602
FT intron
FT /*tag= e
FT /number= 2
FT 96603..96751
FT exon
FT /*tag= f
FT /number= 3
FT 96752..96831
FT intron
FT /*tag= g
FT /number= 3
FT 96832..96915
FT exon
FT /*tag= h

FT
XX
XX WO200268657-A2.
XX
XX 06-SEP-2002.
XX
XX 08-FEB-2002; 2002WO-US003623.
XX
XX 26-FEB-2001; 2001US-0270871P.
XX
XX 26-MAR-2001; 2001US-00816095.
XX
XX (PEXE) PE CORP NY.
XX
XX Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;
XX WPI; 2002-713380/77.
XX P-PSDB; ADI03930.
XX New human enzyme proteins, useful for treating or diagnosing disorders
PT associated with abnormal expression of the protein, in drug screening
PT assays and pharmacogenomic analysis.
XX
XX Claim 4; SEQ ID NO 3; 127pp; English.
XX
XX The invention relates to a novel isolated enzyme protein and encoding
CC polynucleotides. The protein shows a high degree of similarity to a
CC glucuronyltransferase cloned from a rabbit brain cDNA library. The
CC peptides and nucleic acid molecules are useful as models for the
CC development of human therapeutic targets, aid in the identification of
CC therapeutic proteins, and serve as targets for the development of human
CC therapeutic agents. The peptide may be used in drug screening assays, in
CC assays to determine the biological activity of the protein, to raise
CC antibodies or to elicit another immune response, as a reagent in assays
CC designed to quantitatively determine levels of the protein in biological
CC fluids, or as markers for tissues in which the corresponding protein is
CC preferentially expressed. The human enzyme protein is also useful for
CC diagnosing a disease, predisposition to a disease, or treating a disorder
CC characterized by an absence of, inappropriate or unwanted expression of
CC the protein. The antibodies are useful in pharmacogenomic analysis, for
CC inhibiting protein function, or for tissue typing. The nucleic acid
CC molecules are useful as probes, primers, chemical intermediates, or in
CC encoding genomic DNA.
XX
SQ Sequence 99916 BP; 27736 A; 18701 C; 19334 G; 29032 T; 0 U; 5113 Other;
Query Match 50.2%; Score 25.6; DB 6; Length 99916;
Best Local Similarity 70.8%; Pred. No. 40;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAG 49
Db 1819 CAGACCCGCGGAGAGCAAGGGGTCCTCCGCGCTCTCTTCGCAG 1866
RESULT 11
ADI097587
ID ADQ97587 standard; DNA; 116297 BP.
XX
XX ADQ97587;
XX
XX 07-OCT-2004 (first entry)
XX
XX Human cancer associated sequence HD10-009, SEQ ID 564.
XX
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX Homo sapiens.
XX WO2004060304-A2.
XX
XX 22-JUL-2004.
XX


```
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 3070; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful for
XX preventing, treating or ameliorating medical conditions e.g. by protein
XX or gene therapy. The genes are isolated from a range of human tissues
XX disclosed in the specification. The nucleic acids, proteins, antibodies
XX and (ant)agonists are useful in the diagnosis, treatment and prevention
XX of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung, or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 331 BP; 87 A; 83 C; 78 G; 83 T; 0 U; 0 Other;
XX
XX Query Match 47.8%; Score 24.4; DB 4; Length 331;
XX Best Local Similarity 73.8%; Pred. NO. 32;
XX Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 2 CAGAACTGGCAAGGGCAAGGTCACCTAGGGCCCTGTGC 43
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 244 CAGGACTGACACAGGGGTACAGGCCACTGTGGGAATATC 285
```

```
RESULT 15
AAL36704
ID AAL36704 standard; DNA; 331 BP.
XX
XX AC AAL36704;
XX
XX DT 08-JAN-2002 (first entry)
XX
XX DE Human musculoskeletal system related polynucleotide SEQ ID NO 3069.
XX
XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200155367-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001338.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 14-JUL-2000; 2000US-0217496P.
XX 26-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220363P.
XX 14-AUG-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225466P.
XX 14-AUG-2000; 2000US-0225467P.
XX 14-AUG-2000; 2000US-0225477P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 23-AUG-2000; 2000US-0227182P.
XX 30-AUG-2000; 2000US-0227009P.
XX 01-SEP-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0228927P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
```

PR 08-SEP-2000; 2000US-02311413P.
PR 08-SEP-2000; 2000US-02311414P.
PR 08-SEP-2000; 2000US-02320801P.
PR 08-SEP-2000; 2000US-02320801P.
PR 12-SEP-2000; 2000US-02311568P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451937/48.
DR Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
PT cancers and also for testing and detection e.g. diagnosis.
PT
XX Example 2; SEQ ID NO 3069; 781pp + Sequence Listing; English.
PS The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 331 BP; 87 A; 83 C; 78 G; 83 T; 0 U; 0 Other;

Query Match 47.8%; Score 24.4; DB 4; Length 331;

Best Local Similarity 73.8%; Pred. No. 32;

Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGGCGCTCTGTC 43

DB 244 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 285

Search completed: February 7, 2006, 12:49:05
Job time : 34.0376 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:27:59 ; Search time 275.097 Seconds
(without alignments)
10538.142 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acsgaactgcgaagagcca.....agggcctctgtcaccaggga 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	98.0	187383	14 AC138954	AC138954 Homo sapi
C 2	50	98.0	192943	8 AC012645	AC012645 Homo sapi
C 3	29.6	58.0	249737	14 AC094359	AC094359 Rattus no
C 4	29.6	58.0	251990	14 AC128457	AC128457 Rattus no
5	27.6	54.1	189212	14 AC109240	AC109240 Mus muscu
6	27.6	54.1	230186	9 AL731678	AL731678 Mouse DNA
7	27.6	54.1	239692	14 AC158791	AC158791 Mus muscu
C 8	27.6	54.1	242194	14 AC091314	AC091314 Mus muscu
9	27.6	54.1	325223	14 AC109504	AC109504 Mus muscu
10	27.4	53.7	173029	9 AC132271	AC132271 Mus muscu
11	27.4	53.7	198890	9 AC131679	AC131679 Mus muscu
C 12	26.6	52.2	231980	14 AC094305	AC094305 Rattus no
C 13	26.6	52.2	261198	14 AC106528	AC106528 Rattus no
C 14	26.6	52.2	282320	14 AC149647	AC149647 Bos tauru
15	26.2	51.4	136581	8 AC008459	AC008459 Homo sapi
C 16	26.2	51.4	137324	14 AC010459	AC010459 Homo sapi
17	26.2	51.4	143976	14 AC141543	AC141543 Rattus no
C 18	26.2	51.4	184015	9 AC125458	AC125458 Mus muscu

19	26.2	51.4	252996	14 AC130153	AC130153 Rattus no
20	26.2	51.4	345997	1 BX569692	BX569692 Synechoco
21	26	51.0	87746	8 AC109826	AC109826 Homo sapi
22	25.8	50.6	49004	8 AL160288	AL160288 Human DNA
C 23	25.8	50.6	56702	8 AY923052	AY923052 Homo sapi
24	25.8	50.6	150236	14 AC023292	AC023292 Homo sapi
25	25.6	50.2	689	8 HSA339409	HSA339409 Homo sapi
26	25.6	50.2	703	8 HSA339162	HSA339162 Homo sapi
27	25.6	50.2	720	8 HSA324773	HSA324773 Homo sapi
C 28	25.6	50.2	92019	8 AL450320	AL450320 Human DNA
29	25.6	50.2	99916	6 AR438683	AR438683 Sequence
30	25.6	50.2	99916	6 AX925386	AX925386 Sequence
C 31	25.6	50.2	179691	4 BX296515	BX296515 Pig DNA s
32	25.6	50.2	184512	8 AC079804	AC079804 Homo sapi
C 33	25.6	50.2	191318	14 AC020670	AC020670 Homo sapi
C 34	25.6	50.2	191937	8 AC100823	AC100823 Homo sapi
C 35	25.4	49.8	153028	9 AL928739	AL928739 Mouse DNA
C 36	25.4	49.8	186647	9 AC129302	AC129302 Mus muscu
C 37	25.4	49.8	187065	9 AC113014	AC113014 Mus muscu
C 38	25.4	49.8	190780	9 AC153508	AC153508 Mus muscu
C 39	25	49.0	142943	8 AC099405	AC099405 Papio anu
C 40	25	49.0	166513	14 AC073787	AC073787 Mus muscu
41	25	49.0	180223	9 AL589692	AL589692 Mouse DNA
C 42	25	49.0	201676	14 AC105367	AC105367 Sus scrof
43	25	49.0	208651	9 AC090493	AC090493 Genomic s
C 44	25	49.0	209182	9 AC149052	AC149052 Mus muscu
45	25	49.0	216022	14 AC156179	AC156179 Bos tauru

ALIGNMENTS

RESULT 1
AC138954
LOCUS
DEFINITION
Homo sapiens chromosome 16 clone RP11-828J20, WORKING DRAFT
SEQUENCE, 5 unordered pieces.
187383 bp DNA linear HTG 21-JAN-2003

AC138954.1 GI:27805366
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
AC138954
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 (bases 1 to 187383)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 187383)
DOE Joint Genome Institute.
Direct Submission
Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1612692
Center clone name: RPCI-11_828J20

Summary Statistics
Consensus quality: 181858 bases at least Q40
Consensus quality: 184513 bases at least Q30
Consensus quality: 185456 bases at least Q20
Estimated insert size: 18000; agarose-fp estimation
Estimated insert size: 186983; sum-of-contigs estimation
Quality coverage: 8.77 in Q20 bases; agarose-fp estimation
Quality coverage: 8.44 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1225: contig of 1225 bp in length
 * 1226 1325: gap of unknown length
 * 1326 34247: contig of 32922 bp in length
 * 34248 71258: contig of 36911 bp in length
 * 71259 117359: gap of unknown length
 * 117360 117555: contig of 46196 bp in length
 * 117556 187393: gap of unknown length
 * 187394 69729: contig of 69729 bp in length.

FEATURES

source

1. .187383
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-828J20"
 /clone_lib="RP11 human BAC library 11"
 1226. .1325
 /estimated_length=unknown
 34248. .34347
 /estimated_length=unknown
 71259. .71358
 /estimated_length=unknown
 117555. .117654
 /estimated_length=unknown

ORIGIN

Query Match 98.0%; Score 50; DB 14; Length 187383;
 Best Local Similarity 100.0%; Pred. No. 5.2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 50
 |||||
 Db 41734 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 41783

RESULT 2

AC012645/c 192943 bp DNA linear PRI 07-JAN-2004
 LOCUS Homo sapiens chromosome 16 clone RP11-455F5, complete sequence.
 AC012645
 AC012645.7 GI:40714622
 HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

1 (bases 1 to 192943)
 DOE Joint Genome Institute, Stanford Human Genome Center and Los
 Alamos National Laboratory.
 Direct Submission

TITLE

Unpublished

DOE Joint Genome Institute.

DOE Joint Genome Institute.

Direct Submission

Submitted (31-OCT-1998) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

3 (bases 1 to 192943)

DOE Joint Genome Institute.

Direct Submission

Submitted (06-WAR-2002) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

4 (bases 1 to 192943)

Stanford Human Genome Center and Los Alamos National Laboratory.

DOE Joint Genome Institute

Direct Submission

TITLE

JOURNAL

Submitted (07-JAN-2004) DOE Joint Genome Institute, 2800 Mitchell

Drive, Walnut Creek, CA 94598, USA

On Jan 7, 2004 this sequence version replaced gi:19172834.

Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center and Los Alamos

National Laboratory

www.sbgsc.stanford.edu

Quality: Phrap Quality >=40 100% of Sequence;

Estimated Total Number of Errors is 0.

FEATURES

source

Location/Qualifiers
 1. .192943
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-455F5"

ORIGIN

Query Match 98.0%; Score 50; DB 8; Length 192943;
 Best Local Similarity 100.0%; Pred. No. 5.2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 50
 |||||
 Db 53010 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 52961

RESULT 3

AC094359/c

LOCUS

Rattus norvegicus clone CH230-3C19, WORKING DRAFT SEQUENCE, 3

DEFINITION

unordered pieces.

AC094359.6 GI:30467240

HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 249737)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Caesar, H., Center, A.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Y., Chen, Z.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Y., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guetara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshew, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaokemele, O., Okwuonu, G., Olampunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Soderstrom, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 249737)
Worley, K.C.

Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 249737)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24942531.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GALI
Center clone name: CH230-3C19
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 236591 bases at least Q40
Consensus quality: 239010 bases at least Q30
Consensus quality: 241077 bases at least Q20
Estimated insert size: 247522; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 89596: contig of 89596 bp in length
* 89597 89696: gap of unknown length

* 89697 248439: contig of 158743 bp in length
* 248440 248539: gap of unknown length
* 248540 249737: contig of 1198 bp in length.

FEATURES
source
1..249737
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3C19"
54592..54643
/note="clone boundary
clone_end:SP6
site:EcoRI
end_sequence:BH307413"
89597..89696
/estimated_length=unknown
89697..90719
/note="wgs contig"
248440..248539
/estimated_length=unknown

misc_feature
gap
misc_feature
gap

ORIGIN
Query Match 58.0%; Score 29.6; DB 14; Length 249737;
Best Local Similarity 79.5%; Pred. No. 0.98;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCATCTAGGCGCTCTGTCA 44
|||||
Db 222041 ACAGATCTGGCAAGAGGCGAGGTCAGTCTGCTGCTCTTA 221998
|||||

RESULT 4
AC128457/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-85E19, *** SEQUENCING IN PROGRESS
*** 9 unordered pieces.
AC128457
AC128457.2 GI:23267595
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
VERSION
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 251990)
Muzny, D., Maric, Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hollands, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Napua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Pastermak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinkle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS

Unpublished
2 (bases 1 to 251990)

TITLE JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 251990)
Rat Genome Sequencing Consortium.

REFERENCE AUTHORS TITLES JOURNAL

Submitted (10-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GYVR
Center clone name: CH230-85E19

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 190117 bases at least Q40

Consensus quality: 195766 bases at least Q30

Consensus quality: 199840 bases at least Q20

Estimated insert size: 189592; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will

* be preserved.
* 1 11689: contig of 11689 bp in length
* 11690 11789: gap of unknown length
* 11790 23645: contig of 224856 bp in length
* 23646 236745: gap of unknown length
* 236746 241468: contig of 4723 bp in length
* 241469 241568: gap of unknown length
* 241569 242996: contig of 1428 bp in length
* 242997 243096: gap of unknown length
* 243097 244281: contig of 1185 bp in length
* 244282 244381: gap of unknown length
* 244382 245514: contig of 1133 bp in length
* 245515 246732: contig of 1118 bp in length
* 246733 246832: gap of unknown length
* 246833 248947: contig of 2115 bp in length
* 248948 251990: contig of 2943 bp in length.
* 249048 Location/Qualifiers
1..251990
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-85E19"
1..1408
/note="wgs end extension
clone_end:Sp6"
complement(6981..7869)
/note="clone boundary
clone_end:Sp6
site:ECORI
end sequence:BH304395"
11689..11789
/estimated_length=unknown
11790..12959
/note="wgs_contig"
13461..14737
/note="wgs_contig"
33176..34795
/note="wgs_contig"
37663..39153
/note="wgs_contig"
67847..69250
/note="wgs_contig"
72977..74360
/note="wgs_contig"
131426..132574
/note="wgs_contig"
236646..236745
/estimated_length=unknown
241469..241568
/estimated_length=unknown
242997..243096
/estimated_length=unknown
244282..244381
/estimated_length=unknown
245515..245614
/estimated_length=unknown
246733..246832
/estimated_length=unknown
248948..249047
/estimated_length=unknown
ORIGIN
Query Match 58.0%; Score 29.6; DB 14; Length 251990;
Best Local Similarity 79.5%; Pred.No.0.98;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ACAGACTGGCAAGAGGCAAGAGTCACTGAGGCGCTCTGTCA 44
DB 207842 ACAGATCTGGCAAGAGGCGGAGGTGACTGCTGCTCTTA 207799

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-39H12 is from the RPCI-23 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

FEATURES

source

```
1. .230186
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="X"
   /clone="RP23-39H12"
   /clone_lib="RPCI-23"
```

ORIGIN

```
Query Match      54.1%; Score 27.6; DB 9; Length 230186;
Best Local Similarity 78.6%; Pred. No. 6.3;
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      8  TGGCAAGAGCGCAAGGTCACCTGAGGGCTCTGTCCACCAG 49
      |||||
Db      129419 TGGCCACACGCAACGTAAGTCTGAGGGCTCTGTCCACCAG 129460
      |||||
```

RESULT 7

AC158791

LOCUS

```
DEFINITION      Mus musculus chromosome 1 clone RP23-247C6 map 1, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
```

AC158791

AC158791.3 GI:62461205

HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.

Mus musculus (house mouse)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 239692)

Birren, B., Nusbaum, C. and Lander, E.

Mus musculus chromosome 1, clone RP23-247C6

Unpublished

2 (bases 1 to 239692)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (23-MAR-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 239692)

REFERENCE

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Nicol, R., Norbu, C., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (10-APR-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Apr 10, 2005 this sequence version replaced gi:62198378.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L31561

Center clone name: 247_C_6

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*	88410:	contig of 88410 bp in length
*	88411	88510: gap of unknown length
*	88511	92994: contig of 4484 bp in length
*	93094:	gap of unknown length
*	93095	110663: contig of 17569 bp in length
*	110664	110763: gap of unknown length
*	110764	137281: contig of 26518 bp in length
*	137282	137381: gap of unknown length
*	137382	154681: contig of 17300 bp in length
*	154682	154781: gap of unknown length
*	154782	157222: contig of 2441 bp in length
*	157223	157322: gap of unknown length
*	157323	167805: contig of 10483 bp in length
*	167806	167905: gap of unknown length
*	167906	239692: contig of 71787 bp in length.

Location/Qualifiers

source

1. .239692

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-247C6"
/clone_lib="RPCI-23 Female Mouse BAC"
88411..88510
/estimated_length=unknown
92995..93094
/estimated_length=unknown
110664..110763
/estimated_length=unknown
137282..137381
/estimated_length=unknown
154682..154781
/estimated_length=unknown
157223..157322
/estimated_length=unknown
167806..167905
/estimated_length=unknown

ORIGIN
Query Match      54.1%; Score 27.6; DB 14; Length 239692;
Best Local Similarity 72.0%; Pred. No. 6.3;
Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAGAACTGGCAAGAGCGAGAGGTCACCTGAGGGCTCTGTCCACCAGG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227642 ACAGTGCAGCGCAGAGGCTGAGAGTCCAGAGGTCCCTTCACACAGG 227691

RESULT 8
AC091314/c
LOCUS
DEFINITION Mus musculus clone RP23-57B2, WORKING DRAFT SEQUENCE, 17 unordered
            pieces.
AC091314
VERSION AC091314.3 GI:22325114
KEYWORDS HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 242194)
            Birren,B., Nusbaum,C. and Lander,B.
            Mus musculus, clone RP23-57B2
            Unpublished
REFERENCE 2 (bases 1 to 242194)
            Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., Lander,E., Larocque,K.,
            Lanazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G.,
            Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McSwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Sougnuez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (13-APR-2001) Whitehead Institute/MIT Center for Genome

```

REFERENCE AUTHORS

```

Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 242194)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 20, 2002 this sequence version replaced gi:14336555.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13330
Center clone name: 57 B.2
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 237992 bases at least Q40
Consensus quality: 239546 bases at least Q30
Consensus quality: 240147 bases at least Q20
Insert size: 230000; agarose-fp
Insert size: 240594; sum-of-contigs
Quality coverage: 13.7 in Q20 bases; agarose-fp
Quality coverage: 13.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
*
1 22307: contig of 22307 bp in length
* 22308 22407: gap of 100 bp
* 22408 23603: contig of 1196 bp in length
* 23604 23703: gap of 100 bp
* 23704 26091: contig of 2388 bp in length
* 26092 26191: gap of 100 bp
* 26192 28560: contig of 2369 bp in length
* 28561 28660: gap of 100 bp
* 28661 31163: contig of 2503 bp in length
* 31164 31263: gap of 100 bp
* 31264 32354: contig of 1091 bp in length
* 32355 32454: gap of 100 bp
* 32455 34755: contig of 2301 bp in length
* 34756 34855: gap of 100 bp
* 34856 37785: contig of 2930 bp in length
* 37786 37885: gap of 100 bp
* 37886 42925: contig of 5040 bp in length
* 42926 43025: gap of 100 bp

```

TITLE JOURNAL

```

* 43026 47226: contig of 4201 bp in length
* 47227 47326: gap of 100 bp
* 47327 57802: contig of 10476 bp in length
* 57803 72870: gap of 100 bp
* 72871 72970: contig of 14968 bp in length
* 72971 72971: gap of 100 bp
* 89541 89540: contig of 16570 bp in length
* 89541 89540: gap of 100 bp
* 89541 162712: contig of 73072 bp in length
* 162713 162812: gap of 100 bp
* 162813 185801: contig of 22989 bp in length
* 185802 185901: gap of 100 bp
* 185902 220592: contig of 34691 bp in length
* 220593 220593: gap of 100 bp
* 220693 242194: contig of 21502 bp in length.
FEATURES
    source
        1..242194
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone_lib="RPC1-23 Female Mouse BAC"
            1..22307
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:left
                22308..22407
                    /estimated_length=100
                22408..23603
                    /note="assembly_fragment"
                23604..23703
                    /estimated_length=100
                23704..26091
                    /note="assembly_fragment"
                26092..26191
                    /estimated_length=100
                26192..28560
                    /note="assembly_fragment"
                28561..28660
                    /estimated_length=100
                28661..31163
                    /note="assembly_fragment"
                31164..31263
                    /estimated_length=100
                31264..32354
                    /note="assembly_fragment"
                32355..32454
                    /estimated_length=100
                32455..34755
                    /note="assembly_fragment"
                34756..34855
                    /estimated_length=100
                34856..37785
                    /note="assembly_fragment"
                37786..37885
                    /estimated_length=100
                37886..42925
                    /note="assembly_fragment"
                42926..43025
                    /estimated_length=100
                43026..47226
                    /note="assembly_fragment"
                47227..47326
                    /estimated_length=100
                47327..57802
                    /note="assembly_fragment"
                57803..57902
                    /estimated_length=100
                57903..72870
                    /note="assembly_fragment"
                72871..72970
                    /estimated_length=100
                72971..89540

```

```

/note="assembly_fragment"
89541..89640
/estimated_length=100
89641..162712
/note="assembly_fragment"
162713..162812
/estimated_length=100
162813..185801
/note="assembly_fragment"
185802..185901
/estimated_length=100
185902..220592
/note="assembly_fragment"
220593..220692
/estimated_length=100

Query Match      54.1%; Score 27.6; DB 14; Length 242194;
Best Local Similarity 78.6%; Pred. No. 6.3;
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      8 TGGCAAGAGGCAAGAGTCACTGAGGGCCCTGTGTACCCAG 49
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB     106371 TGGCAACAGGCCAACAGTAAGTACTGAGGCCCTGTGTCCACG 106330

RESULT 9
AC109504      329223 bp      DNA      linear      HTG 13-MAR-2004
LOCUS      Mus musculus chromosome 13 clone RP23-184C21 map 13, WORKING DRAFT
DEFINITION      SEQUENCE, 20 unordered pieces.
ACCESSION      AC109504
VERSION      AC109504.5 GI:45430118
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus
            Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 329223)
Birren,B., Nussbaum,C. and Lander,E.
Mus musculus chromosome 13, clone RP23-184C21
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 329223)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhaltier,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Lander,E., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Maddison,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (04-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 329223)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhaltier,B., Camarata,J., Chang,J., Choepeil,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,

```

Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 13, 2004 this sequence version replaced gi:20389429.
All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL COMMENT

FEATURES source

1. .329223
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="13"
/map="13"
/clone="RP23-184C21"
/clone_lib="RPCI-23 Female Mouse BAC"

Location/Qualifiers

1. .1913
/notes="assembly_fragment"
1914. .2013
/estimated_length=100
2014. .27036
/notes="assembly_fragment"
27037. .27136
/estimated_length=100
27137. .31366
/notes="assembly_fragment"
31367. .31466
/estimated_length=100
31467. .85408
/notes="assembly_fragment"
85409. .85508
/estimated_length=100
85509. .89621
/notes="assembly_fragment"
89622. .89721
/estimated_length=100
89722. .104564
/notes="assembly_fragment"
104565. .104664
/estimated_length=100
104665. .112440
/notes="assembly_fragment"
112441. .112540
/estimated_length=100
112541. .131698
/notes="assembly_fragment"
131699. .131798
/estimated_length=100
131799. .142284
/notes="assembly_fragment"
142285. .142384
/estimated_length=100
142385. .151005
/notes="assembly_fragment"
151006. .151105
/estimated_length=100
151106. .157224
/notes="assembly_fragment"
157225. .157324
/estimated_length=100
157325. .166896
/notes="assembly_fragment"
166897. .166996

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIFR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L20868

Center clone name: 184 C.21

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 192122 bases at least Q40

Consensus quality: 194367 bases at least Q30

Consensus quality: 195058 bases at least Q20

Insert size: 182000; agarose-gel

Insert size: 195349; sum-of-contigs

Quality coverage: 7.4 in Q20 bases; agarose-gel

Quality coverage: 6.9 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 20 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1913: contig of 1913 bp in length

1914 2013: gap of 100 bp

2014 27036: contig of 25023 bp in length

27037 31366: gap of 100 bp

31367 31366: contig of 4230 bp in length

31467 31466: gap of 100 bp

31467 85408: contig of 53942 bp in length

85409 85508: gap of 100 bp

85509 89621: contig of 4113 bp in length

89622 89721: gap of 100 bp

89722 104564: contig of 14843 bp in length

104565 104664: gap of 100 bp

104665 112440: contig of 7776 bp in length

112441 112540: gap of 100 bp

112541 131698: contig of 19158 bp in length

131699 131798: gap of 100 bp

131799 142284: contig of 10486 bp in length

142285 142384: gap of 100 bp

142385 151005: contig of 8621 bp in length

151006 151105: gap of 100 bp

151106 157224: contig of 6119 bp in length

157225 157324: gap of 100 bp

157325 166896: gap of 100 bp

166897 166996: gap of 100 bp

166997 166996: gap of 100 bp

```
misc_feature      /estimated length=100
166997..222799
/notes="assembly_fragment"
gap              /estimated_length=100
222800..222899
misc_feature      /notes="assembly_fragment"
222900..229569
gap              229570..229669

Query Match      54.1%; Score 27.6; DB 14; Length 329223;
Best Local Similarity 72.0%; Pred. No. 6.4;
Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ACAGAACTGCGAAGAGGCAAGAGGTCACTGAGGGCTCTGTCCACCCAGG 50
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199292 ACAGTCCGCGCAGAGGCTGGAGGTCACAGAGTCCCTTTTCACACAGG 199341

RESULT 10
AC132271
LOCUS      AC132271      173029 bp      DNA      linear      ROD 08-MAR-2005
DEFINITION Mus musculus BAC clone RP24-338E23 from 5, complete sequence.
ACCESSION  AC132271
VERSION    AC132271.5 GI:60301731
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 173029)
AUTHORS   Tomlinson, C., Haglund, K. and Bielicki, L.
TITLE     The sequence of Mus musculus BAC clone RP24-338E23
JOURNAL   Unpublished (2001)
REFERENCE  2 (bases 1 to 173029)
AUTHORS   McPherson, J.D. and Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  3 (bases 1 to 173029)
AUTHORS   Wilson, R.K.
TITLE     Direct Submission
JOURNAL   Submitted (09-FEB-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  4 (bases 1 to 173029)
AUTHORS   Wilson, R.K.
TITLE     Direct Submission
JOURNAL   Submitted (26-FEB-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE  5 (bases 1 to 173029)
AUTHORS   Wilson, R.K.
TITLE     Direct Submission
JOURNAL   Submitted (08-MAR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT    On Feb 26, 2005 this sequence version replaced gi:58801841.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: M_BB0338E23

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
```

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source

1..173029

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="5"

/map="5"

/clone_lib="RPCI-24"

/clone="RP24-338E23"

24291..24314

/note="Sequence derived from PCR product of project DNA."

24956..25048

/note="Sequence derived from one plasmid subclone."

45800..45813

/note="Sequence derived from one plasmid subclone."

139522..139794

/note="Sequence derived from PCR product of genomic DNA."

139840..140750

/note="Sequence derived from PCR product of genomic DNA."

140751..141175

/note="Unresolved simple sequence repeat."

141176..141456

/note="Sequence derived from PCR product of genomic DNA."

163787..163801

/note="Sequence derived from one plasmid subclone."

ORIGIN

Query Match 53.7%; Score 27.4; DB 9; Length 173029;

Best Local Similarity 75.6%; Pred. No. 7.5;

Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCCACCCAGGA 51

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 167276 CTGCCAAAAGCAACAGCTCACTGAGAGCCTCTGCCTCAAGGA 167320

RESULT 11

LOCUS

AC131679

DEFINITION

Mus musculus BAC clone RP23-339I6 from chromosome 5, complete

sequence.

AC131679

VERSION

AC131679.2 GI:40949631

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 198890)

AUTHORS

Haakenson, W., Haglund, K., Bielicki, L. and Meyer, R.

TITLE

The sequence of Mus musculus BAC clone RP23-339I6

JOURNAL

Unpublished (2001)

REFERENCE

2 (bases 1 to 198890)

AUTHORS

Wilson, R.

TITLE

Sequencing of Mus musculus

JOURNAL

Unpublished (2001)

```
REFERENCE
AUTHORS
TITLE
JOURNAL
3 (bases 1 to 198890)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 198890)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (29-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 198890)
Wilson,R.K.
Direct Submission
Submitted (16-JAN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
6 (bases 1 to 198890)
Wilson,R.
Direct Submission
Submitted (06-FEB-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 16, 2004 this sequence version replaced gi:22475611.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----
Center project name: M_BA0339106
-----
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tatenio in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

```
FEATURES
Source
1..198890
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-33916"
/clone_lib="RPCI-23"
47..120
/rpt_family="ID"
repeat_region
297..460
/rpt_family="L2"
repeat_region
```

```
repeat_region
1033..1208
/rpt_family="B4"
/rpt_family="MaLR"
1462..1571
/rpt_family="MIR"
2980..3174
/rpt_family="L1"
3779..3963
/rpt_family="B2"
4280..5114
/rpt_family="L1"
5301..5581
/rpt_family="B4"
5886..6053
/rpt_family="MaLR"
6168..6368
/rpt_family="B2"
7206..7332
/rpt_family="Alu"
7999..8016
/notes="Sequence derived from one plasmid subclone."
8227..8357
/rpt_family="B4"
8720..8930
/rpt_family="B4"
8984..9066
/rpt_family="B4"
9912..10042
/rpt_family="B4"
10327..10743
/rpt_family="ERVK"
11314..11597
/rpt_family="B4"
11804..11912
/rpt_family="L1"
11930..12128
/rpt_family="RMER15"
12129..12320
/rpt_family="B2"
12321..12677
/rpt_family="RMER15"
12692..12840
/rpt_family="L1"
12990..13061
/rpt_family="MaLR"
13062..13245
/rpt_family="B2"
13246..13573
/rpt_family="MaLR"
13636..13829
/rpt_family="B2"
15079..15326
/rpt_family="MaLR"
16411..16566
/rpt_family="B4"
17194..17300
/rpt_family="Alu"
17404..17610
/rpt_family="MER1_type"
18519..18954
/rpt_family="MaLR"
21064..21238
/rpt_family="B2"
21800..21952
/rpt_family="Alu"
22082..22205
/rpt_family="Alu"
23272..23353
/rpt_family="Alu"
23446..23514
/rpt_family="Alu"
23665..23827
/rpt_family="L2"
repeat_region
```

```

repeat_region 25366..25417
/rpt_family="MER1_type"
repeat_region 25459..25611
/rpt_family="Alu"
repeat_region 27083..27314
/rpt_family="L1"
repeat_region 28449..28648
/rpt_family="B2"
repeat_region 28678..28866
/rpt_family="L1"
repeat_region 29380..29467
/rpt_family="Alu"
repeat_region 29938..30007
/rpt_family="ID"
repeat_region 30078..30353
/rpt_family="B4"
repeat_region 31735..31881
/rpt_family="B4"
repeat_region 31927..32015
/rpt_family="B4"
repeat_region 32016..32220
/rpt_family="B2"
repeat_region 32221..32322
/rpt_family="B4"
repeat_region 34803..34851
/rpt_family="Alu"
repeat_region 35801..35859
/rpt_family="ERV1"
repeat_region 36514..36689
/rpt_family="MER1_type"
repeat_region 36727..36998
/rpt_family="B4"
repeat_region 38298..38494
/rpt_family="ERVK"
repeat_region 38675..38730

Query Match 53.7%; Score 27.4; DB 9; Length 198890;
Best Local Similarity 75.6%; Pred. No. 7.5;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 CTGGCAAGAGGCGAAGAGTCACTGAGGGCTCTGTCAACCCAGGA 51
|||||
Db 11491 CTGCCAAAAGCAACAGCTCACTGAGAGGCTCTGCTCAAGGA 11535

RESULT 12
AC094305/c
LOCUS AC094305 231980 bp DNA linear HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-3P23, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION AC094305
VERSION AC094305.11 GI:30467414
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,W.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Deigado,O., Benson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Poster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Iozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umami,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,R., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 231980)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 231980)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24952938.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAIZ
Center clone name: CH230-3P23
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 205998 bases at least Q40

```


AUTHORS
TITLE
JOURNAL

Worley, K.C.
Direct Submission
Submitted (19-JUN-2004) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 282320)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

On Jun 29, 2005 this sequence version replaced gi:50540557.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FDTZ
Center clone name: CH240-68K2
----- Summary Statistics

Assembly program: Atlas 3.0:
Consensus quality: 233356 bases at least Q40
Consensus quality: 235130 bases at least Q30
Consensus quality: 236805 bases at least Q20
Estimated insert size: 239331; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 22 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence.
as soon as it is available and the accession number will
be preserved.

1
7906: contig of 7906 bp in length
7907
7956: gap of 50 bp
7957
21416: contig of 13460 bp in length
21417
21466: gap of 50 bp
32425: contig of 10959 bp in length
32426
32475: gap of 50 bp
32476
71121: contig of 38646 bp in length
71122
71171: gap of 50 bp
73347: contig of 2176 bp in length
73348
107554: gap of 34207 bp
123340: contig of 15786 bp in length
123341
123390: gap of 50 bp
127911: contig of 4521 bp in length
127912
127961: gap of 50 bp
127962
135626: contig of 7665 bp in length
135627
135676: gap of 50 bp
144612: contig of 8936 bp in length
144613
144662: gap of 50 bp
145739: contig of 1077 bp in length
145740
151919: gap of 6180 bp
158564: contig of 6645 bp in length
158565
158614: gap of 50 bp
158615
182943: contig of 24329 bp in length

182944
183044
184051: contig of 1008 bp in length
184052
184151: gap of unknown length
184152
185427: contig of 1276 bp in length
185428
185527: gap of unknown length
185528
185596: contig of 1069 bp in length
185597
186596: gap of unknown length
186597
188035: contig of 1339 bp in length
188036
188135: gap of unknown length
188136
189300: contig of 1165 bp in length
189301
189400: gap of unknown length
189401
191191: contig of 1791 bp in length
191192
191291: gap of unknown length
191292
198350: contig of 6959 bp in length
198351
203718: contig of 5368 bp in length
203719
203818: gap of unknown length
203819
209340: contig of 5522 bp in length
209341
209440: gap of unknown length
209441
282320: contig of 72880 bp in length.

FEATURES
source

1. 282320
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-68K2"
7907. 7956
/estimated_length=50
21417. 21466
/estimated_length=50
32426. 32475
/estimated_length=50
71122. 71171
/estimated_length=50
73348. 107554
/estimated_length=34207
123341. 123390
/estimated_length=50
127912. 127961
/estimated_length=50
135627. 135676
/estimated_length=50
144613. 144662
/estimated_length=50
145740. 151919
/estimated_length=6180
158565. 158614
/estimated_length=50
182944. 183043
/estimated_length=unknown
184052. 184151
/estimated_length=unknown
185428. 185527
/estimated_length=unknown
186597. 186696
/estimated_length=unknown
188036. 188135
/estimated_length=unknown
189301. 189400
/estimated_length=unknown
191192. 191291
/estimated_length=unknown
198251. 198350

Query Match 52.2%; Score 26.6; DB 14; Length 282320;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACGTGAGGGCCCTCTGTCCACCAGG 50
|||||
DB 45334 CACACGTGGCAAGAGGCAAGAGGTCATTTAGGTTCCCTGAAGCCACAGG 45382
|||||

RESULT 15
AC008459 136581 bp DNA linear PRI 04-FEB-2003
LOCUS Homo sapiens chromosome 5 clone CTC-352D11, complete sequence.
DEFINITION AC008459
ACCESSION AC008459
VERSION AC008459.7 GI:28201472
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 136581)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 136581)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 136581)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 136581)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Feb 4, 2003 this sequence version replaced gi:14579681.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1
NOTE: Shatter libraries failed to verify the dinucleotide repeat
region 3558-5794. Unsure number of repeat copies 3558-5794.
Forced join at 4399.
FEATURES
source
1. .136581
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-352D11"
3558..5794
misc_feature
/note="NOTE: Shatter libraries failed to verify the
dinucleotide repeat region 3558-5794. Unsure number of
repeat copies 3558-5794. Forced join at 4399."
ORIGIN
Query Match 51.4%; Score 26.2; DB 8; Length 136581;
Best Local Similarity 79.5%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 7 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCTAC 45
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19816 CTGGCAGATAAGCAGGAGGCCCTCTGGAGGCTCTGTCTAC 19854

Search completed: February 7, 2006, 15:20:44
Job time : 281.097 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:30:54 ; Search time 10.8591 Seconds
(without alignments)
8348.349 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236

Perfect score: 51

Sequence: 1 acgaactggaagagga.....aggcctctgtcaccaggga 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25.6	50.2	99916	3	US-09-816-095-3
2	24.4	47.8	601	3	US-09-949-016-122934
3	24.4	47.8	601	3	US-09-949-016-123028
4	24.4	47.8	601	3	US-09-949-016-123122
5	24.4	47.8	601	3	US-09-949-016-123216
6	24.4	47.8	601	3	US-09-949-016-123310
7	24.4	47.8	1023	3	US-09-902-540-9591
8	24.4	47.8	13346	3	US-09-902-540-1089
9	24.4	47.8	32379	3	US-09-949-016-15217
10	24.4	47.8	32379	3	US-09-949-016-15218
11	24.4	47.8	32379	3	US-09-949-016-15219
12	24.4	47.8	32379	3	US-09-949-016-15220
13	24.4	47.8	32379	3	US-09-949-016-15221
14	24.4	47.8	32379	3	US-09-949-016-15222
15	24.4	47.8	32379	3	US-09-949-016-15223
16	24.4	47.8	32379	3	US-09-949-016-15224
17	24.4	47.8	32379	3	US-09-949-016-15225
18	24.4	47.8	32379	3	US-09-949-016-15226
19	24.4	47.8	64377	3	US-09-949-016-15212
20	24.4	47.8	64377	3	US-09-949-016-15213
21	24.4	47.8	64377	3	US-09-949-016-15214
22	24.4	47.8	64377	3	US-09-949-016-15215
23	24.4	47.8	64377	3	US-09-949-016-15216
24	23.2	45.5	3089	3	US-10-104-047-914

25	22.8	44.7	1191	3	US-09-902-540-8380	Sequence 8380, Ap
c 26	22.8	44.7	2022	3	US-09-023-655-416	Sequence 416, App
c 27	22.8	44.7	2088	3	US-09-949-016-4237	Sequence 4237, Ap
c 28	22.8	44.7	2737	3	US-09-902-540-3798	Sequence 3798, Ap
c 29	22.8	44.7	4231	3	US-09-949-016-809	Sequence 809, App
c 30	22.8	44.7	7035	3	US-09-902-540-878	Sequence 878, App
c 31	22.8	44.7	18551	3	US-09-902-540-1187	Sequence 1187, Ap
c 32	22.8	44.7	21593	3	US-09-949-016-15979	Sequence 15979, A
c 33	22.8	44.7	96878	3	US-09-949-016-12551	Sequence 12551, A
c 34	22.8	44.7	114793	3	US-10-148-806-3	Sequence 187024, A
c 35	22.6	44.3	601	3	US-09-949-016-187025	Sequence 187025, A
c 36	22.6	44.3	601	3	US-09-949-016-187026	Sequence 187026, A
c 37	22.6	44.3	601	3	US-09-949-016-187025	Sequence 187025, A
c 38	22.6	44.3	56616	3	US-09-949-016-17085	Sequence 17085, A
c 39	22.6	44.3	56616	3	US-09-949-016-15402	Sequence 15402, A
c 40	22.6	44.3	152070	3	US-09-949-016-120678	Sequence 120678, A
c 41	22.4	43.9	601	3	US-09-949-016-158390	Sequence 158390, A
c 42	22.4	43.9	601	3	US-09-949-016-158391	Sequence 158391, A
c 43	22.4	43.9	601	3	US-09-949-016-158392	Sequence 158392, A
c 44	22.4	43.9	601	3	US-09-949-016-158392	Sequence 158392, A
c 45	22.4	43.9	601	3	US-09-949-016-158393	Sequence 158393, A

ALIGNMENTS

RESULT 1

US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816, 095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(99916)
; OTHER INFORMATION: n = A, T, C or G
US-09-816-095-3

Query Match	50.2%	Score	25.6;	DB	3;	Length	99916;
Best Local Similarity	70.8%	Pred. No.	10;				
Matches	34;	Conservative	0;	Mismatches	14;	Indels	0;
Gaps	0;						
Qy	2	CAGACTGGCAAGAGCAAGGTCAC	TGAGGGCTCTGTCCACCGAG	49			
Db	1819	CAGACCGCGGAGAGCCAGCGG	TCCCTCCGCTCTCTCGCAG	1866			

RESULT 2

US-09-949-016-122934
; Sequence 122934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122934

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 3
US-09-949-016-123028
; Sequence 123028, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123028
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123028

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 4
US-09-949-016-123122
; Sequence 123122, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123122
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123122

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 5
US-09-949-016-123216
; Sequence 123216, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123216
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123216

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 6
US-09-949-016-123310
; Sequence 123310, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123310
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123310
```

Query Match 47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGGGCCTCTGTC 43
DB 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 7

US-09-902-540-9591/c
; Sequence 9591, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCES: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9591
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9591

Query Match 47.8%; Score 24.4; DB 3; Length 1023;
Best Local Similarity 68.0%; Pred. No. 11;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGGGCCTCTGTCACCCAGGA 51
DB 866 CAGAACTGGCGCGCCAGCAGCAGTCACTCAGGGCCTCCAACTCCAGGA 817

RESULT 8

US-09-902-540-1089/c
; Sequence 1089, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCES: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1089
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1089

Query Match 47.8%; Score 24.4; DB 3; Length 1346;
Best Local Similarity 68.0%; Pred. No. 19;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGGGCCTCTGTCACCCAGGA 51
DB 4227 CAGAACTGGCGCGCCAGCAGCAGTCACTCAGGGCCTCCAACTCCAGGA 4178

RESULT 9

US-09-949-016-15217/c
; Sequence 15217, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCES: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15217
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15217

Query Match 47.8%; Score 24.4; DB 3; Length 32379;
Best Local Similarity 73.8%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGGGCCTCTGTC 43
DB 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289

RESULT 10

US-09-949-016-15218/c
; Sequence 15218, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCES: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15218
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15218

Query Match 47.8%; Score 24.4; DB 3; Length 32379;
Best Local Similarity 73.8%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGGGCCTCTGTC 43
DB 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289

RESULT 11

US-09-949-016-15219/c
; Sequence 15219, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15219
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15219

Query Match
Best Local Similarity 47.8%; Score 24.4; DB 3; Length 32379;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289

RESULT 12
US-09-949-016-15220/c
; Sequence 15220, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15220
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15220

Query Match
Best Local Similarity 47.8%; Score 24.4; DB 3; Length 32379;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289

RESULT 13
US-09-949-016-15221/c
; Sequence 15221, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15221
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15221

Query Match
Best Local Similarity 47.8%; Score 24.4; DB 3; Length 32379;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289

RESULT 14
US-09-949-016-15222/c
; Sequence 15222, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15222
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15222

Query Match
Best Local Similarity 47.8%; Score 24.4; DB 3; Length 32379;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289

RESULT 15
US-09-949-016-15223/c
; Sequence 15223, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15223
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15223

Query Match      47.8%; Score 24.4; DB 3; Length 32379;
Best Local Similarity 73.8%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy      2 CAGAACTGGCAAGAGAGAGGTCACAGAGGCGCTCTGTC 43
      |||||
Db      1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 1289
      |||||

```

Search completed: February 7, 2006, 15:27:25
Job time : 12.8591 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 13:52:58 ; Search time 22.0694 Seconds
(without alignments)
1936.698 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236

Perfect score: 51

Sequence: 1 acgaactgcgaagagca.....aggctctgtcaccaggga 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24.2	47.5	1656	7 US-10-750-185-34830	Sequence 34830, A
2	24.2	47.5	1656	7 US-10-750-623-34830	Sequence 34830, A
3	23.6	46.3	600	7 US-10-750-185-20700	Sequence 20700, A
4	23.6	46.3	600	7 US-10-750-623-20700	Sequence 20700, A
5	23.6	46.3	903	7 US-10-750-185-40559	Sequence 40559, A
6	23.6	46.3	903	7 US-10-750-623-40559	Sequence 40559, A
7	23.4	45.9	201	8 US-11-124-368A-3178	Sequence 3178, Ap
8	23.2	45.5	34875	7 US-10-775-169-316	Sequence 316, App
9	23.2	45.5	164810	8 US-11-121-086-6	Sequence 6, Appl
10	23.2	45.5	172543	8 US-11-121-086-6	Sequence 6, Appl
11	23	45.1	944	7 US-10-750-185-64443	Sequence 64443, A
12	23	45.1	944	7 US-10-750-623-64443	Sequence 64443, A
13	22.8	44.7	201	7 US-10-995-561-84242	Sequence 84242, A
14	22.8	44.7	40000	7 US-10-995-561-13512	Sequence 13512, A
15	22.8	44.7	182303	8 US-11-121-086-45	Sequence 45, Appl
16	22.4	43.9	1144	7 US-10-750-185-25067	Sequence 25067, A
17	22.4	43.9	1144	7 US-10-750-623-25067	Sequence 25067, A
18	22.4	43.9	2077	7 US-10-750-185-44085	Sequence 44085, A
19	22.4	43.9	2077	7 US-10-750-623-44085	Sequence 44085, A
20	22.4	43.9	2686	7 US-10-750-185-35306	Sequence 35306, A
21	22.4	43.9	2686	7 US-10-750-623-35306	Sequence 35306, A
22	22.4	43.9	4643	8 US-11-136-527-3270	Sequence 3270, Ap

RESULT 1

US-10-750-185-34830
; Sequence 34830, Application US/10750185
; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 34830

; LENGTH: 1656

; TYPE: DNA

; ORGANISM: Bovine 19866880752334

US-10-750-185-34830

Query Match 47.5%; Score 24.2; DB 7; Length 1656;

Best Local Similarity 71.1%; Pred. No. 7.9;

Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 5 AACTGCCAAGAGCAAGGTCACTCAGGGCGCTCTCTCACCAG 49

Db 578 AACAGTCAGTGAGANTAGAGATCGAAGAGAGCCCTCTCCACCCAG 622

RESULT 2

US-10-750-623-34830

; Sequence 34830, Application US/10750623

; Publication No. US20050287531A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

```

; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20700
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine
; ORGANISM: Bovine 19866880752334
US-10-750-623-34830

Query Match 47.5%; Score 24.2; DB 7; Length 1656;
Best Local Similarity 71.1%; Pred. No. 7.9;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 5 AACTGGCAAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAG 49
|||||
DB 578 AACAGTCAGTGAATAGATCGAAGAGAGGCTCTGCCACCCAG 622

RESULT 3
US-10-750-185-20700
; Sequence 20700, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20700
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine
; ORGANISM: Bovine MMBT20675
US-10-750-185-20700

Query Match 46.3%; Score 23.6; DB 7; Length 600;
Best Local Similarity 76.3%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCCT 38
|||||
DB 559 ATAAACCTGGCAAGGCAAGAGGACCCCTGAGGTCT 596

RESULT 4
US-10-750-623-20700
; Sequence 20700, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31

```

```

; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20700
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine
; ORGANISM: Bovine MMBT20675
US-10-750-623-20700

Query Match 46.3%; Score 23.6; DB 7; Length 600;
Best Local Similarity 76.3%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCCT 38
|||||
DB 559 ATAAACCTGGCAAGGCAAGAGGACCCCTGAGGTCT 596

RESULT 5
US-10-750-185-40559
; Sequence 40559, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40559
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Bovine
; ORGANISM: Bovine 19866881097008
US-10-750-185-40559

Query Match 46.3%; Score 23.6; DB 7; Length 903;
Best Local Similarity 76.3%; Pred. No. 12;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCCT 38
|||||
DB 780 ATAAACCTGGCAAGGCAAGAGGACCCCTGAGGTCT 817

RESULT 6
US-10-750-623-40559
; Sequence 40559, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31

```

```

; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40559
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Bovine 19866881097008
US-10-750-623-40559

Query Match      46.3%; Score 23.6; DB 7; Length 903;
Best Local Similarity 76.3%; Pred. No. 12; Mismatches 0; Gaps 0;
Matches 29; Conservative 0; Indels 9;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCT 38
    |||||
DB 780 ATAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCT 817
    |||||

RESULT 7
US-11-124-368A-3178/c
; Sequence 3178, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: C0001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3178
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-3178

Query Match      45.9%; Score 23.4; DB 8; Length 201;
Best Local Similarity 64.7%; Pred. No. 10; Mismatches 17; Indels 0; Gaps 0;
Matches 33; Conservative 1;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTGTCACCCAGGA 51
    |||||
DB 120 AAACAATTTTCAGGGAGTCRTGAAGCCCTTCTGGGCGCTCTTCCACCCAGGA 70
    |||||

RESULT 8
US-10-775-169-316/c
; Sequence 316, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 34875
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-775-169-316

Query Match      45.5%; Score 23.2; DB 7; Length 34875;
Best Local Similarity 70.5%; Pred. No. 35; Mismatches 13; Indels 0; Gaps 0;
Matches 31; Conservative 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTGTCACCCAGG 50
    |||||
DB 12441 CTGGCACAGAGACACACGGGGCACAGGAGGGCGTCAGCCACGCAGG 12398
    |||||

RESULT 9
US-11-121-086-4/c
; Sequence 4, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 164810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-4

Query Match      45.5%; Score 23.2; DB 8; Length 164810;
Best Local Similarity 70.5%; Pred. No. 48; Mismatches 13; Indels 0; Gaps 0;
Matches 31; Conservative 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTGTCACCCAGG 50
    |||||
DB 8707 CTGGCACAGAGACACACGGGGCACAGGAGGGCGTCAGCCACGCAGG 8664
    |||||

RESULT 10
US-11-121-086-6/c
; Sequence 6, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 172543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-6

Query Match      45.5%; Score 23.2; DB 8; Length 172543;
Best Local Similarity 70.5%; Pred. No. 48; Mismatches 13; Indels 0; Gaps 0;
Matches 31; Conservative 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTGTCACCCAGG 50
    |||||
DB 143325 CTGGCACAGAGACACACGGGGCACAGGAGGGCGTCAGCCACGCAGG 143282
    |||||

RESULT 11
US-10-750-185-64443/c
```

```
; Sequence 64443, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64443
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Bovine 19866881015183
; US-10-750-185-64443

Query Match 45.1%; Score 23; DB 7; Length 944;
Best Local Similarity 68.1%; Pred. No. 20;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ACAGAACTGCGAAAGAGGCAAGAGGTCACTGAGGGCTCTGTCACCC 47
Db 78 AGAGCAATGGAATGAAGCCAGAGGACAGTACAGACCTGTTTCACCC 32

RESULT 12
US-10-750-623-64443/c
; Sequence 64443, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64443
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Bovine 19866881015183
; US-10-750-623-64443

Query Match 45.1%; Score 23; DB 7; Length 944;
Best Local Similarity 68.1%; Pred. No. 20;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ACAGAACTGCGAAAGAGGCAAGAGGTCACTGAGGGCTCTGTCACCC 47
Db 78 AGAGCAATGGAATGAAGCCAGAGGACAGTACAGACCTGTTTCACCC 32

RESULT 13
US-10-995-561-84242
; Sequence 84242, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
```

```
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84242
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-84242

Query Match 44.7%; Score 22.8; DB 7; Length 201;
Best Local Similarity 68.2%; Pred. No. 17;
Matches 30; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

QY 6 ACTGCGAAAGAGGCAAGAGGTCACTGAGGGCTCTGTCACCCAG 49
Db 69 ACCAGCTAGGAGGCCAAGACACTGATGGCACTTTCCCCAG 112

RESULT 14
US-10-995-561-13512
; Sequence 13512, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13512
; LENGTH: 40000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(40000)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-2
; US-10-995-561-13512

Query Match 44.7%; Score 22.8; DB 7; Length 40000;
Best Local Similarity 68.2%; Pred. No. 50;
Matches 30; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

QY 6 ACTGCGAAAGAGGCAAGAGGTCACTGAGGGCTCTGTCACCCAG 49
Db 36556 ACCAGCTAGGAGGCCAAGACACTGATGGCACTTTCCCCAG 36599

RESULT 15
US-11-121-086-45/c
; Sequence 45, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 45
```

```

; LENGTH: 182303
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-45

Query Match      44.7%; Score 22.8; DB 8; Length 182303;
Best Local Similarity 66.0%; Pred. No. 68;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY      2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCACCACAGGA 51
      ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db      145227 CAGAACATGGAGGAGTGTGAGAGGAGGAGGGCTTCGGTGTCCCATGA 145178

```

Search completed: February 7, 2006, 14:11:06
Job time : 22.0694 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:27:59 ; Search time 9908.9 Seconds
(without alignments)
10538.142 Million cell updates/sec

Title: US-10-623-108-5

Perfect score: 1837

Sequence: 1 gagagtgagatggcggcg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_to.*

10: gb_sts.*

11: gb_ey.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1725	93.9	1813	8	BC013992 Homo sapi
2	1708.4	93.0	1866	6	AR164078 Sequence
3	1708.4	93.0	1866	6	BD058089 Antisense
4	1708.4	93.0	1866	6	CS063502 Sequence
5	1708.4	93.0	1866	6	CS081283 Sequence
6	1708.4	93.0	1866	6	AX770530 Sequence
7	1708.4	93.0	1866	8	HSERK1
8	1708.4	93.0	3732	6	AX676347 Sequence
9	1632	88.8	1745	8	HUMKER1A
10	1546.8	84.2	2567	8	AK091009
11	1540.8	83.9	2722	8	HSM806004
12	1201	65.4	1284	8	HS44KDP
13	1140	62.1	1140	11	AY888192
14	1140	62.1	1140	11	AY890458
15	1140	62.1	1140	11	AY893270
16	1138.4	62.0	1140	6	AX722274
17	1138.4	62.0	1140	11	AY335595
18	1138.4	62.0	1140	11	AY890764

19	1138.4	62.0	1140	11	AY890765
20	1138.4	62.0	1896	6	A84423
21	1138.4	62.0	1896	6	BD082524
22	1138.4	62.0	1896	6	BD228337
23	1138.4	62.0	1896	6	Q0829947
24	1138.4	62.0	1896	6	AR279902
25	1138.4	62.0	1896	6	AX427322
26	1080.2	58.8	1125	6	AX803349
27	1061	57.8	1783	9	SC46779
28	1060.4	57.7	1875	9	RATERK1A
29	1049.6	57.1	1736	9	RNU12008
30	1048	57.0	1721	9	MMERK1
31	1033.4	56.3	1718	9	RATERK1
32	1030.2	56.1	1715	9	AR016531
33	1025	55.8	1747	6	AR036244
34	1025	55.8	1747	6	AR171290
35	1025	55.8	1747	6	AR172791
36	1025	55.8	1747	6	I34199
37	1025	55.8	1747	6	I34199
38	1015.4	55.3	1074	8	AY033607
39	965.4	52.6	1238	9	RNM4PKIN
40	958.6	52.2	1875	6	A84441
41	958.6	52.2	1875	6	BD082533
42	958.6	52.2	1875	6	Q0829965
43	958.6	52.2	1875	6	AR279911
44	958.6	52.2	1875	6	AX427340
45	939.6	51.1	1632	9	BC013754

ALIGNMENTS

RESULT 1	BC013992	1813 bp	mRNA	linear	PRI 29-JUN-2004
LOCUS	BC013992				
DEFINITION	Homo sapiens mitogen-activated protein kinase 3, mRNA (cDNA clone MGC:20180 IMAGE:3634492), complete cds.				
ACCESSION	BC013992				
VERSION	BC013992.1				GI:15559270
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1813)				
AUTHORS	Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Bickton, Y.S., Krzywicki, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1813)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer				

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granice, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 28 Row: d Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers

1..1813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20180 IMAGE:3634492"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH MGC 8"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1..1813
/gene="MAPK3"
/notes="synonyms: P44ERK1, P44MAPK, ERK1, PRKM3"
/db_xref="GeneID:5595"
/db_xref="MIM:601795"
12..1151
/gene="MAPK3"
/codon_start=1
/product="mitogen-activated protein kinase 3"
/protein_id="AAH13992.1"
/db_xref="GI:15559271"
/db_xref="GeneID:5595"
/db_xref="MIM:601795"
/translation="MAAARQQGGGEPRTGVRGVRGVEVMVGQPPDVPRYTQ
LQVIGGAYGMVSSAYDHYRKTVAIKKISPFPHQTYCORTLRBTQILLRFRHENVIG
LRDILRASLEARDYVIQDLMETDLYKLSQSLNDHIQVFLYQILRLKYLHSA
NVLRDLKPSNLLINTTDLKICDFGLARIADPEHDHTGLFYVATRWYRAPEIMLN
SKGYTKSIDVSGCICLWLNRPFPCKHVLQDLNHLILGILSPSQEDLNCIINMK
ARNYLSKSTKVMAKLFPPKSDSKALDLRLMTFNPKNRITVEALAHPLYEQY
DPTDEPVASEPFTFAMELDDLPKRLKELIFQTARFPQGVLEAP"

ORIGIN

Query Match 93.9%; Score 1725; DB 8; Length 1813;
Best Local Similarity 97.2%; Pred. No. 8.6e-293;
Matches 1786; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 GAGGAGTGGAGATGGCGGCGCGCGCTCAGGGGGGCGGGGCGGGAGCCCCGTAGAA 60
DB 1 GAGGAGTGGAGATGGCGGCGCGCGCTCAGGGGGGCGGGGCGGGAGCCCCGTAGAA 60
QY 61 CCAGGGGGTCCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCCTTCG 120
DB 61 CCAGGGGGTCCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCCTTCG 120
QY 121 ACCTGGGCGCGCTACACGACAGTTCAGTACATCGCGGAGGCGGCTACGGCATGGTCA 180

DB 121 ACCTGGGCGCGCTACACGACAGTTCAGTACATCGCGGAGGCGGCTACGGCATGGTCA 180
QY 181 GCTCGGCGCTATGACCAAGTGGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 240
DB 181 GCTCGGCGCTATGACCAAGTGGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 240
QY 241 AACATCAGACCTACTGCGCAGGCGCAGCTCCGGGAGATCCAGATCTCTGCTGCGCTTCCGCG 300
DB 241 AACATCAGACCTACTGCGCAGGCGCAGCTCCGGGAGATCCAGATCTCTGCTGCGCTTCCGCG 300
QY 301 ATGAGAATGTCTATCGGCATCCGAGACATTTCTCGGGGCGTCCACCTTGGGAAGCCATGAGAG 360
DB 301 ATGAGAATGTCTATCGGCATCCGAGACATTTCTCGGGGCGTCCACCTTGGGAAGCCATGAGAG 360
QY 361 ATGTCTACATTTGTGAGGACCTGATGGAGACTGACTGTGTAAGTTGCTGTAAGAGCCAGC 420
DB 361 ATGTCTACATTTGTGAGGACCTGATGGAGACTGACTGTGTAAGTTGCTGTAAGAGCCAGC 420
QY 421 AGCTGAGCAATCACCATATCTCTACTTCTCTACCGAGATCCTGGGGGCGCTCAAGTACA 480
DB 421 AGCTGAGCAATCACCATATCTCTACTTCTCTACCGAGATCCTGGGGGCGCTCAAGTACA 480
QY 481 TCCACTCCGCCCAACGTGCTCCACCGAGATCTAAAGCCCTTCAAACCTGCTCATCAACACCA 540
DB 481 TCCACTCCGCCCAACGTGCTCCACCGAGATCTAAAGCCCTTCAAACCTGCTCATCAACACCA 540
QY 541 CTTGGGACCTTAAGATTTGTGATTTTGGCGCTGGCGCGGATTCGGATCTCTGAGCATGACC 600
DB 541 CTTGGGACCTTAAGATTTGTGATTTTGGCGCTGGCGCGGATTCGGATCTCTGAGCATGACC 600
QY 601 ACACGGGCTTCTGACGGAGTATGTGGCTACCGCTGCTGTTACCGGGGCCCGAGAGATCATGC 660
DB 601 ACACGGGCTTCTGACGGAGTATGTGGCTACCGCTGCTGTTACCGGGGCCCGAGAGATCATGC 660
QY 661 TGAATCTCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720
DB 661 TGAATCTCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720
QY 721 AGATGCTCTCTAAACGGGCCCATTTCCCTGGCAAGCATCTACCTGGATCAGCTCAACCCACA 780
DB 721 AGATGCTCTCTAAACGGGCCCATTTCCCTGGCAAGCATCTACCTGGATCAGCTCAACCCACA 780
QY 781 TTCTGGGCACTCTGGGCTCCCATCCAGAGGACCTGAATTTGTATCATCAACATGAAGG 840
DB 781 TTCTGGGCACTCTGGGCTCCCATCCAGAGGACCTGAATTTGTATCATCAACATGAAGG 840
QY 841 CCGGAAACTACCTACAGTCTCTGCGCTCCAGACCAAGGTGGCTTGGGCGCAAGCTTTTCC 900
DB 841 CCGGAAACTACCTACAGTCTCTGCGCTCCAGACCAAGGTGGCTTGGGCGCAAGCTTTTCC 900
QY 901 CCAAGTCAGACTCCAAAGCCCTTGACCTGCGCGGATGTTAACTTTAAACCCCAATA 960
DB 901 CCAAGTCAGACTCCAAAGCCCTTGACCTGCGCGGATGTTAACTTTAAACCCCAATA 960
QY 961 AACGATCACAGTGGAGGAAGCGCTGGCTCACCCCTTACCTGGAGCAGTACTATGACCGGA 1020
DB 961 AACGATCACAGTGGAGGAAGCGCTGGCTCACCCCTTACCTGGAGCAGTACTATGACCGGA 1020
QY 1021 CGGATGAGCCAGTGGCGGAGGAGCCCTTCACTTGGCCATGGAGCTGGATGATACCTTA 1080
DB 1021 CGGATGAGCCAGTGGCGGAGGAGCCCTTCACTTGGCCATGGAGCTGGATGATACCTTA 1080
QY 1081 AGGAGGGGCTGAAGAGCTCATCTTCCAGGAGACGAGCTTCCAGCCCGGAGTGTCTCG 1140
DB 1081 AGGAGGGGCTGAAGAGCTCATCTTCCAGGAGACGAGCTTCCAGCCCGGAGTGTCTCG 1140
QY 1141 AGGCCCCCTAGCCAGACAGACATCTCTGACCTTGGGGCTTGAACAGAACTGGGAAAG 1200
DB 1141 AGGCCCCCTAGCCAGACAGACATCTCTGACCTTGGGGCTTGAACAGAACTGGGAAAG 1200
QY 1201 AGGCAAGAGGTCACTGAGGGGCTCTGTACCCAGGAGACTTGTCTGCTGCTGCCCTTCTCC 1260
DB 1201 AGGCAAGAGGTCACTGAGGGGCTCTGTACCCAGGAGACTTGTCTGCTGCTGCCCTTCTCC 1260

[illegible]

[illegible]


```
QY 723 ATGCTCTCTAAGCGGCCCATCTTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAAT 782
Db 784 ATGCTCTCTAAGCGGCCCATCTTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAAT 843
QY 783 CTGGGATCTCTGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAGGCC 842
Db 844 CTGGGATCTCTGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAGGCC 903
QY 843 CGAACTTACCTACAGTCTCTGCGCTCCCAAGCAAGGTGCTGGGCCAAGCTTTTCCCC 902
Db 904 CGAACTTACCTACAGTCTCTGCGCTCCCAAGCAAGGTGCTGGGCCAAGCTTTTCCCC 963
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTAACCTTTAAACCCCAATAA 962
Db 964 AAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTAACCTTTAAACCCCAATAA 1023
QY 963 CGATCACAGTGAAGGAGCGCTGCTCACCCTTACTGGGACGACTATGACCCGAG 1022
Db 1024 CGATCACAGTGAAGGAGCGCTGCTCACCCTTACTGGGACGACTATGACCCGAG 1083
QY 1023 GATGAGCACTGCGGAGGAGCCCTTCACTCTGCGCATGAGCTGGATGACCTACCTAAG 1082
Db 1084 GATGAGCACTGCGGAGGAGCCCTTCACTCTGCGCATGAGCTGGATGACCTACCTAAG 1143
QY 1083 GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACGACGCTTCCAGCCCGGAGTGTGGAG 1142
Db 1144 GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACGACGCTTCCAGCCCGGAGTGTGGAG 1203
QY 1143 GCCCCTTAGCCCAAGACATCTCTGCACTGCGGAGCGCTTCCAGCCCGGAGTGTGGAG 1202
Db 1204 GCCCCTTAGCCCAAGACATCTCTGCACTGCGGAGCGCTTCCAGCCCGGAGTGTGGAG 1243
QY 1203 GCAAGAGTCACTGAGGCGCTCTGTACCCAGAGACCTGCGCTCTGCTGCGCCCTCTCCCG 1262
Db 1244 -----GACCTGCTCTCTGCTGCGCCCTCTCTCCCG 1272
QY 1263 CCAGACTGTTAGAAAATGGACACTGTGCCCCAGCCGACCTTGGCAGCCAGCCCGGGGT 1322
Db 1273 CCAGACTGTTAGAAAATGGACACTGTGCCCCAGCCGACCTTGGCAGCCAGCCCGGGGT 1332
QY 1323 GAGAGATGGGCTGGCCACCTCTCTCTTTGCTGAGGCGCTCCAGCTTCAGGCGAGCCAG 1382
Db 1333 GAGAGATGGGCTGGCCACCTCTCTCTTTGCTGAGGCGCTCCAGCTTCAGGCGAGCCAG 1392
QY 1383 GCCTTCTCTCCACCCGCGCTCCCAAGCGGCGCTCGGAGCTCAGGTGGCCCAAGTTC 1442
Db 1393 GCCTTCTCTCCACCCGCGCTCCCAAGCGGCGCTCGGAGCTCAGGTGGCCCAAGTTC 1452
QY 1443 AATCTCCC---GCTGCTGCTGCTGCGCCCTTTACCTTCCCAGCGTCCCAAGTCTCTGGCAG 1499
Db 1453 AATCTCCGCTGCTGCTGCTGCGCCCTTTACCTTCCCAGCGTCCCAAGTCTCTGGCAG 1512
QY 1500 TTCTGGAATGGAAGGTTCTGGTGGCCCACTCTGCTGGAAGGGGAGAGGTGGAGGTGGG 1559
Db 1513 TTCTGGAATGGAAGGTTCTGGTGGCCCACTCTGCTGGAAGGGGAGAGGTGGAGGTGGG 1572
QY 1560 GGGCGCTGAGTAGGACTCAGGGCCATGCGTCCCGCCCTCATCTCATTTCAAAACCCACCC 1619
Db 1573 GGGCGCTGAGTAGGACTCAGGGCCATGCGTCCCGCCCTCATCTCATTTCAAAACCCACCC 1632
QY 1620 TAGTTTCCCTGAAGGAATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1679
Db 1633 TAGTTTCCCTGAAGGAATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1692
QY 1680 GGGCCGAATCCCTCTCTGTCGCTCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1739
Db 1693 GGGCCGAATCCCTCTCTGTCGCTCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
QY 1740 GGTGAGCAGAAGTGGAGCTGGGGGGCTGGAGAGCCCGCGCGCTGCGCTGCGCTGCGCTGAC 1799
Db 1753 GGTGAGCAGAAGTGGAGCTGGGGGGCTGGAGAGCCCGCGCGCTGCGCTGCGCTGCGCTGAC 1812
```

```
QY 1800 CCGTCTATATATAATATAGAGATGCTGCTATGGCTG 1837
Db 1813 CCGTCTATATATAATATAGAGATGCTGCTATGGCTG 1850

RESULT 6
LOCUS AX770530 1866 bp DNA linear PAT 02-JUL-2003
DEFINITION Sequence 36 from Patent WO03031650.
ACCESSION AX770530
VERSION AX770530.1 GI:32437906
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Munnes, M., Gehrman, M., Wick, M. and Schmitz, G.
TITLE Genes and proteins for prevention, prediction, prognosis and
therapy of cardiovascular disease
JOURNAL Patent: WO 03031650-A 36 17-APR-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1. .1866
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 93.0%; Score 1708.4; DB 6; Length 1866;
Best Local Similarity 97.0%; Pred. No. 7e-290;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;

QY 3 GAGTGTGAGATGCGCGCGCGCTCAGGGGGGGGGGGGGGGGAGAGCCCGTAGAAC 62
Db 64 GAGTGTGAGATGCGCGCGCGCGCTCAGGGGGGGGGGGGGGGGAGAGCCCGTAGAAC 123
QY 63 GAGGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 122
Db 124 GAGGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 183
QY 123 GTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 182
Db 184 GTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 243
QY 183 TCGGGCTATGACCACTGCGCGCAAGACTCGCGTGGCCCATCAAGAGATCAGCCCTTCGAA 242
Db 244 TCGGGCTATGACCACTGCGCGCAAGACTCGCGTGGCCCATCAAGAGATCAGCCCTTCGAA 303
QY 243 CATCAGACCTTACTGCGCAGCGCTCCGGGAGATCCAGATCTGTGCGCTTCCGCGCAT 302
Db 304 CATCAGACCTTACTGCGCAGCGCTCCGGGAGATCCAGATCTGTGCGCTTCCGCGCAT 363
QY 303 GAGATGTCTATCGGCTATCGAGACATCTGCGGGGGCTCCACCTTGGAGGCGATGAGAT 362
Db 364 GAGATGTCTATCGGCTATCGAGACATCTGCGGGGGCTCCACCTTGGAGGCGATGAGAT 423
QY 363 GTCTATCTGTCAGGACCTGATGGAGACTGACTGTGTAAGTTGCTGAAAGCCAGCAG 422
Db 424 GTCTATCTGTCAGGACCTGATGGAGACTGACTGTGTAAGTTGCTGAAAGCCAGCAG 483
QY 423 CTGAGCAATGACCATATCTGCTATCTTCTCTACAGATCTCTGCGGGGGCTCAAGTACATC 482
Db 484 CTGAGCAATGACCATATCTGCTATCTTCTCTACAGATCTCTGCGGGGGCTCAAGTACATC 543
QY 483 CATCTCGCCAACTGTCTCCACCGAGATCTAAGGCGCTCCAACTGCTCATCACACACC 542
Db 544 CATCTCGCCAACTGTCTCCACCGAGATCTAAGGCGCTCCAACTGCTCATCACACACC 603
QY 543 TGGAGCCTTAAGATTTGTGATTTTCGGCTGCGCGCGGATTCGCCATCTCTGAGCATGACCAC 602
Db 604 TGGAGCCTTAAGATTTGTGATTTTCGGCTGCGCGCGGATTCGCCATCTCTGAGCATGACCAC 663
```


1489	TGCA	TTCTGGCTGAGATGCTCTCTAAACCGGCCCAATCTTCCCTGGCAAGCACTACCTGGAT	1544
768	CAGCTCA	ACACCAATCTCTGGGCATCTCTGGGCTCCCCATCCAGGAGACCTGGAATTGTATC	827
1549	CAGCTCA	ACACCAATCTCTGGGCATCTCTGGGCTCCCCATCCAGGAGACCTGGAATTGTATC	1608
828	ATCA	ACATGAAGGCCCGAAACTACTACGTCTCTGCCCCTCAAGACCAAGTGGCTTGG	887
1609	ATCA	ACATGAAGGCCCGAAACTACTACGTCTCTGCCCCTCAAGACCAAGTGGCTTGG	1668
888	GCCA	AGCTTTTCCCCTCAAGTCCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAAAC	947
1669	GCCA	AGCTTTTCCCCTCAAGTCCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAAAC	1728
948	TTTAA	CCCCCAATAAAACGATATCAGTGGAGGAAGCGCTGCTCACCCCTACCTGGAGCAG	1007
1729	TTTAA	CCCCCAATAAAACGATATCAGTGGAGGAAGCGCTGCTCACCCCTACCTGGAGCAG	1788
1008	TACTAT	GACCCGAGGATGAGCCATGGCCCGAGGAGGCCCTTCACTTCCGCATGGAGCTG	1067
1789	TACTAT	GACCCGAGGATGAGCCATGGCCCGAGGAGGCCCTTCACTTCCGCATGGAGCTG	1848
1068	GATG	ACTTACTTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAG	1127
1849	GATG	ACTTACTTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAG	1908
1128	CCCG	GAGTGTGGAGGCCCTTAGCCAGACAGACATCTCTGCACCTGGGGCTCGAAC	1187
1909	CCCG	GAGTGTGGAGGCCCTTAGCCAGACAGACATCTCTGCACCTGGGGCT----	1963
1188	AGAA	CTGGCAAGGCAAGAGTCTAGTGAAGGCTCTGTACCCAGGACCTGCTCTCTG	1247
1964	-----	-----GGACCTGCCTCTG	1977
1248	CCTG	CCCCCTCTCCGCCAGACTGTTAGAAATGGACACTGTGCCACCGCGGACCTTGGC	1307
1978	CCTG	CCCCCTCTCCGCCAGACTGTTAGAAATGGACACTGTGCCACCGCGGACCTTGGC	2037
1308	AGCC	AGGCGGGGTGAGACATGGGCGCTGGCCACTCTCTCTTTGCTGAGGGCTCCAGC	1367
2038	AGCC	AGGCGGGGTGAGACATGGGCGCTGGCCACTCTCTCTTTGCTGAGGGCTCCAGC	2097
1368	TTCA	GGCAGGCCAAGGCTTCTCTCTCCCAACCGCCCTTCCCAACGGGGCTCTGGGAGCTC	1427
2098	TTCA	GGCAGGCCAAGGCTTCTCTCTCCCAACCGCCCTTCCCAACGGGGCTCTGGGAGCTC	2157
1428	AGTG	GGCCCAATCTCTCCGCTGCTGTGTGGCGCCCTTACCTTCCCAAGCGCTCC	1487
2158	AGTG	GGCCCAATCTCTCCGCTGCTGTGTGGCGCCCTTACCTTCCCAAGCGCTCC	2217
1488	AGTCT	CTGGCAGTTCTGGAAATGGAAGGTTCTTGCTGCCCAACCTGCTGAAGGGCAGAG	1547
2218	AGTCT	CTGGCAGTTCTGGAAATGGAAGGTTCTTGCTGCCCAACCTGCTGAAGGGCAGAG	2277
1548	GTG	AGGGGTGGGGCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCCCTCATCTCAT	1607
2278	GTG	AGGGGTGGGGCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCCCTCATCTCAT	2337
1608	CAAA	CCCCCAGCTAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTG	1667
2338	CAAA	CCCCCAGCTAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTG	2397
1668	AGG	AGCAGCGGCGGCGGAATCCCTCTCTGTCAAAGCTGTCACTTCGCGTCCCTCGCT	1727
2398	AGG	AGCAGCGGCGGCGGAATCCCTCTCTGTCAAAGCTGTCACTTCGCGTCCCTCGCT	2457
1728	GCTT	CTGTGTGTGTGAGCAAGTGGAGCTGGGGGGCGTGGAGAGCCCGCGCCCCCTGC	1787
2458	GCTT	CTGTGTGTGTGAGCAAGTGGAGCTGGGGGGCGTGGAGAGCCCGCGCCCCCTGC	2517
1788	CAC	CTCCCTGACCTCTAATATATAAATATAGAGATGTGTCTATGGCTG	1837
2518	CAC	CTCCCTGACCTCTAATATATAAATATAGAGATGTGTCTATGGCTG	2567

Db 948 TCCGGTCCCTCAGCTCGSCCTATGACACAGTGGCGCAAGACTCGGTGGCCATCAAGAG 1007
QY 228 ATCAGCCCTTGGAAATCAGACCTACTGCGACGCGACGCTCGGGAGATCCAGATCCGT 287
Db 1008 ATCAGCCCTTGGAAATCAGACCTACTGCGACGCGACGCTCGGGAGATCCAGATCCGT 1067
QY 288 CTGCGCTTCCGCGCATGAGAAATGTCATCGGCATCCGAGACATTTCTGGGCGCTCCACCCCTG 347
Db 1068 CTGCGCTTCCGCGCATGAGAAATGTCATCGGCATCCGAGACATTTCTGGGCGCTCCACCCCTG 1127
QY 348 GAAGCCATGAGAGATGTCATTTGTGTGAGGACCTGATGGAGACTGACCTGTACAAAGTTG 407
Db 1128 GAAGCCATGAGAGATGTCATTTGTGTGAGGACCTGATGGAGACTGACCTGTACAAAGTTG 1187
QY 408 CTGAAAGCCAGAGCTGAGCAATGACCAATATCTGCTACTTCTTACACAGATCTCGGG 467
Db 1188 CTGAAAGCCAGAGCTGAGCAATGACCAATATCTGCTACTTCTTACACAGATCTCGGG 1247
QY 468 GGCCTCAAGTACATCCACTCCGCAACGCTGCTCCACGAGATCTAAAGCCCTCCCAACCTG 527
Db 1248 GGCCTCAAGTACATCCACTCCGCAACGCTGCTCCACGAGATCTAAAGCCCTCCCAACCTG 1307
QY 528 CTATCAACACCACTCGCAACCTTAAAGTTTGTGATTTCTGGGCTGCGCCCGGATTCGCCAT 587
Db 1308 CTATCAACACCACTCGCAACCTTAAAGTTTGTGATTTCTGGGCTGCGCCCGGATTCGCCAT 1367
QY 588 CTGAGCATGACACACCGGCTTCTGACGAGATATGCGCTACGCGCTGTTACCGGGCC 647
Db 1368 CTGAGCATGACACACCGGCTTCTGACGAGATATGCGCTACGCGCTGTTACCGGGCC 1427
QY 648 CCAGAGATCATGCTGAACCTCCAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGC 707
Db 1428 CCAGAGATCATGCTGAACCTCCAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGC 1487
QY 708 TGCAATTTCTGGCTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGAT 767
Db 1488 TGCAATTTCTGGCTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGAT 1547
QY 768 CAGCTCAACACATTTCTGGGCTATCTGGGCTGCCATCCAGGAGACCTGAATGTATC 827
Db 1548 CAGCTCAACACATTTCTGGGCTATCTGGGCTGCCATCCAGGAGACCTGAATGTATC 1607
QY 828 ATCAACATGAAGCCCGGAACTACCTACAGTCTCTGCCCTCCAAAGCAAGGCTGGCTGG 887
Db 1608 ATCAACATGAAGCCCGGAACTACCTACAGTCTCTGCCCTCCAAAGCAAGGCTGGCTGG 1667
QY 888 GCCAAGCTTTTCCCAAGTCAGACTCCAAAGCCCTTGACCTGCTGGAACCGGATGTTAAC 947
Db 1668 GCCAAGCTTTTCCCAAGTCAGACTCCAAAGCCCTTGACCTGCTGGAACCGGATGTTAAC 1727
QY 948 TTTAACCCCAATAAACGATCAAGTGGAGGAGAGCGCTGGCTCACCCCTACCTGGAGCAG 1007
Db 1728 TTTAACCCCAATAAACGATCAAGTGGAGGAGAGCGCTGGCTCACCCCTACCTGGAGCAG 1787
QY 1008 TACTATGACCCGAGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCGCCATGGAGCTG 1067
Db 1788 TACTATGACCCGAGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCGCCATGGAGCTG 1847
QY 1068 GATGACTTACCTTAAGGAGGCGCTGAAGGAGCTCATTTCTCAGGAGACAGACGCTTCCAG 1127
Db 1848 GATGACTTACCTTAAGGAGGCGCTGAAGGAGCTCATTTCTCAGGAGACAGACGCTTCCAG 1907
QY 1128 CCGGAGTGTGAGGCGCCCTTAGCCACAGACATCTCTGACCCCTGGGCGCTGGAC 1187
Db 1908 CCGGAGTGTGAGGCGCCCTTAGCCACAGACATCTCTGACCCCTGGGCGCT----- 1962
QY 1188 AGAAGTGGCAAGAGGAGGTCACTGAGGCGCTCTGTCAACCGAGGACCTGCTCCCTG 1247
Db 1963 -----GGAAGTGGCTTCTG 1976
QY 1248 CTTGCCCTCTCCCGCAGACTGTTAGAAAATGACACTGTGCCAGCCGCGACTTGGC 1307

Db 1977 CTTGCCCTCTCCCGCAGACTGTTAGAAAATGGACACTGTGTGCCAGCCCGACCTTGGC 2036
QY 1308 AGCCAGGCGCGGGTGGAGCATATGGCCCTGGCCACCTCTCTCTTTGCTGAGGCTCCAGC 1367
Db 2037 AGCCAGGCGCGGGTGGAGCATATGGCCCTGGCCACCTCTCTCTTTGCTGAGGCTCCAGC 2096
QY 1368 TTACGCGAGGCGCAAGGCTTCTCTCCCAACCCGCGCTCCCAACCGGGGCTCGGAGCTC 1427
Db 2097 TTACGCGAGGCGCAAGGCTTCTCTCTCCCAACCCGCGCTCCCAACCGGGGCTCGGAGCTC 2156
QY 1428 AGTGTGCGCCAGTTCATCTCCCGCTGCTGCTGCTGCGCCCTTACTCTTCCCGAGGCTCC 1487
Db 2157 AGTGTGCGCCAGTTCATCTCCCGCTGCTGCTGCTGCGCCCTTACTCTTCCCGAGGCTCC 2216
QY 1488 AGTGTCTGCGCACTCTGGAATGGAAGGCTTCTGGCTGCGCCCAACCTGTCTGAAGGCGCAG 1547
Db 2217 AGTGTCTGCGCACTCTGGAATGGAAGGCTTCTGGCTGCGCCCAACCTGTCTGAAGGCGCAG 2276
QY 1548 GTGAGGCTGCGGGCGCTGAGTAGGGAATCAGGCGCATATGCTGCCCTCATCTCAT 1607
Db 2277 GTGAGGCTGCGGGCGCTGAGTAGGGAATCAGGCGCATATGCTGCCCTCATCTCAT 2336
QY 1608 CAAACCCCAACCTTCTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTG 1667
Db 2337 CAAACCCCAACCTTCTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTG 2396
QY 1668 AGGAGCCAGGCGCGGCGCAATCCCTCTGCTCAAGGCTGCTCACTTGGGCTGCCCTCGCT 1727
Db 2397 AGGAGCCAGGCGCGGCGCAATCCCTCTGCTCAAGGCTGCTCACTTGGGCTGCCCTCGCT 2456
QY 1728 GCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTGCG 1787
Db 2457 GCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTGCG 2516
QY 1788 CACCTCCCTGAGCCCTTAAT 1831
Db 2517 CACCTCCCTGAGCCCTTAAT 2560

RESULT 12

LOCUS HS44KDALP 1284 bp mRNA linear PRI 18-APR-2005
DEFINITION H.sapiens 44kDa protein kinase related to rat ERK1.
ACCESSION Z11696 S38872
VERSION Z11696.1 GI:23882
KEYWORDS protein kinase.
SOURCE Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 1284)
AUTHORS Gonzalez, F.A., Raden, D.L., Rigby, M.R. and Davis, R.J.
TITLE Heterogeneous expression of four MAP kinase isoforms in human tissues
JOURNAL FEBS Lett. 304 (2-3), 170-178 (1992)

PUBMED 131925
REFERENCE 2 (bases 1 to 1284)
AUTHORS Gonzalez, F.A.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1992) Fernando A Gonzalez, Biochemistry and
Molecular Biology, University of Massachusetts Medical School, 373
Plantation St., Worcester, MA, 01605, USA

COMMENT On May 6, 2005 this sequence version replaced gi:250709.
FEATURES
Location/Qualifiers
1..1284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pBluescript II"
/cell_line="HeLa"
/clone_lib="Lambda Zap II"
1..1068

CDS

before 'ATG' to provide Kozak consensus sequence. Each clone is
clonally isolated and full-length sequence-verified.

FEATURES
source
1. .1140
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="FLH019087.01X"
/lab_host="Escherichia coli DH5alpha Tl resistant"
/note="derived from Homo sapiens first strand cDNA library
from placenta and brain"
1. .1140
/gene="MAPK3"
1. .1140
/gene="MAPK3"
/codon_start=1
/transl_table=11
/product="mitogen-activated protein kinase 3"
/protein_id="AA41139.1"
/db_xref="GI:61355431"
/translation="MAAAAGGGGGGPRRTGVRGVRGVEVMVKGPFDVGPRTQ
LQYIGGAYGMSSADHYHVKRKTVAIKISPEFHQYTCORTLREIQILLRPHENVIG
IRDILRASTLEMDYVYIQDLMETLYKLXSLNDHICFYQLIRGLKYTHSA
NVLHRLKPSNLLINTCDLKI CDFGLRIADPEHDHTGFLTYVATRWYRAPEIMLN
SKVYTKSIDWSVGTCLAEMLNRPFPKGHYLDLNLHILGILGPSORDLNCIINMK
ARNYLOSLSKTKVAWKLPPKSDSKALDLDMLTFNPKRITVVEALAHPLYEQYY
DPTDEFAEPTFFAMELDDLPERLKLILFQETARFPQVLEAP"

ORIGIN

Query Match 62.1%; Score 1140; DB 11; Length 1140;
Best Local Similarity 100.0%; Pred. No. 4.3e-190;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 ATGGCGGCGCGGCTCAGGGGGCGGGGGCGGGGAGCCCGGTAGAACCCGAGGGGGTC 71
DB 1 ATGGCGGCGCGGCTCAGGGGGCGGGGGCGGGGAGCCCGGTAGAACCCGAGGGGGTC 60
QY 72 GGGCGGGGGTCCCGGGGAGTGGAGATGGTCAAGGGGCGAGCGCTTCGACGTTGGGCGCG 131
DB 61 GGGCGGGGGTCCCGGGGAGTGGAGATGGTGAAGGGGCGAGCGCTTCGACGTTGGGCGCG 120
QY 132 CGCTACACGCGAGTTGCAGTACATCGGCGAGGGCGGTACGGCATGTGTGTCAGCTCGGCGTAT 191
DB 121 CGCTACACGCGAGTTGCAGTACATCGGCGAGGGCGGTACGGCATGTGTGTCAGCTCGGCGTAT 180
QY 192 GACCACTGGCGAAGACTCGCTGGGCCATCAAGAAGATCAGCCCTTCGAAATCAGACC 251
DB 181 GACCACTGGCGAAGACTCGCTGGGCCATCAAGAAGATCAGCCCTTCGAAATCAGACC 240
QY 252 TACTGCCAGCGACGCTCCGGGAGATCCAGATCTCTGGCTTCGCGCATGAGATGTC 311
DB 241 TACTGCCAGCGACGCTCCGGGAGATCCAGATCTCTGGCTTCGCGCATGAGATGTC 300
QY 312 ATCGGATCCGAGACATTCCTCGGGGGTCCACCTCGAAGCCATGAGATGTCATCAT 371
DB 301 ATCGGATCCGAGACATTCCTCGGGGGTCCACCTCGAAGCCATGAGATGTCATCAT 360
QY 372 GTCGAGCACTGATGAGACTGACCTGTGTA CAAGTTGCTGAAAAGCCAGGAGCTGAGCAAT 431
DB 361 GTCGAGCACTGATGAGACTGACCTGTGTA CAAGTTGCTGAAAAGCCAGGAGCTGAGCAAT 420
QY 432 GACCATATCTGCTACTCTCTACCAAGATCTCGGGGGCTCAAGTACATCCACTCCGCC 491
DB 421 GACCATATCTGCTACTCTCTACCAAGATCTCGGGGGCTCAAGTACATCCACTCCGCC 480
QY 492 AAGTGTCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACTCCGACCTT 551
DB 481 AAGTGTCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACTCCGACCTT 540
QY 552 AAGATTGTGATTTCGGCCCTGGCCGATTTGCCGATCTTCCGATGACCAACACCGGCTTC 611
DB 541 AAGATTGTGATTTCGGCCCTGGCCGATTTGCCGATCTTCCGATGACCAACACCGGCTTC 600

QY 612 CTGACGAGTAGTATGTGGCTACGCGCTGGTACCGGGCCCCAGAGATCATGCTGAATCCCAAG 671
DB 601 CTGACGAGTAGTATGTGGCTACGCGCTGGTACCGGGCCCCAGAGATCATGCTGAATCCCAAG 660
QY 672 GGCTATACCAAGTCCATCGACATCTGGTCTGGGCTGATTCCTGGCTGAGATGCTCTCT 731
DB 661 GGCTATACCAAGTCCATCGACATCTGGTCTGGGCTGATTCCTGGCTGAGATGCTCTCT 720
QY 732 AACCGGCGCATCTCCCTGGCAAGCACTACCTGGATCAGCTCAACACATTCCTGGGCAATC 791
DB 721 AACCGGCGCATCTCCCTGGCAAGCACTACCTGGATCAGCTCAACACATTCCTGGGCAATC 780
QY 792 CTGGGCTCCCATCCAGGAGACCTGAATTCATCAATCATCAACATGAAGCCCGAAATCTAC 851
DB 781 CTGGGCTCCCATCCAGGAGACCTGAATTCATCAATCATCAACATGAAGCCCGAAATCTAC 840
QY 852 CTACAGTCTCTGCCCTCCAAAGCAAGGTGGCTTGGGCAAGCTTTTCCCTCCCAAGTCAGAC 911
DB 841 CTACAGTCTCTGCCCTCCAAAGCAAGGTGGCTTGGGCAAGCTTTTCCCTCCCAAGTCAGAC 900
QY 912 TCCAAAGCCCTTGACCTGCTGACCGGATGTTAACTTTAACCCCAATAAAGCGGATCACA 971
DB 901 TCCAAAGCCCTTGACCTGCTGACCGGATGTTAACTTTAACCCCAATAAAGCGGATCACA 960
QY 972 GTGGAGGAAGCCCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACGATGAGCCA 1031
DB 961 GTGGAGGAAGCCCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACGATGAGCCA 1020
QY 1032 GTGGCGGAGAGCCCTTACCTTCGCCATGGAGCTGGATGACCTACCTTAAGGAGCGGCTG 1091
DB 1021 GTGGCGGAGAGCCCTTACCTTCGCCATGGAGCTGGATGACCTACCTTAAGGAGCGGCTG 1080
QY 1092 AAGGAGCTCATCTCCAGGAGACAGCAGCTTCCAGCCGGAGTGCTGGAGGCCCCCTTAG 1151
DB 1081 AAGGAGCTCATCTTCAGGAGACAGCAGCTTCCAGCCGGAGTGCTGGAGGCCCCCTTAG 1140
RESULT 14
AY890458 1140 bp mRNA linear SYN 29-MAR-2005
LOCUS
DEFINITION
Synthetic construct Homo sapiens clone FLH135362.01X
mitogen-activated protein kinase 3 (MAPK3) mRNA, complete cds.
ACCESSION
AY890458
VERSION
AY890458.1 GI:61363490
KEYWORDS
Human ORF Project.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 1140)
Hines, L., Rolfs, A., Jepsen, D., Moreira, D., Raphael, J., Kelley, F.,
Shen, B., Halleck, A., Koundinya, M., Hu, Y., Zuo, D., Taycher, E.,
Williamson, J. and Labaer, J.
AUTHORS
Cloning of human full-length CDS in Creator (TM) recombinational
vector system
TITLE
Unpublished
JOURNAL
2 (bases 1 to 1140)
Hines, L., Rolfs, A., Jepsen, D., Moreira, D., Raphael, J., Kelley, F.,
Shen, B., Halleck, A., Koundinya, M., Hu, Y., Zuo, D., Taycher, E.,
Williamson, J. and Labaer, J.
AUTHORS
Direct Submission
TITLE
Submitted (04-JAN-2005) Biological Chemistry and Molecular
JOURNAL
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
Cambridge, MA 02141, USA
COMMENT
This CDS clone is a part of a collection of human full-length
expression clones generated by Harvard Institute of Proteomics.
This ORF clone has been cloned with normalized stop-codon. The CDS
has been directionally cloned using BD In-Fusion(TM) cloning system
between the SalI and HindIII sites of the pBIR-Dual vector.
Additional sequences in the clone: 'ACC' after SalI site and
before 'ATG' to provide Kozak consensus sequence. Each clone is
clonally isolated and full-length sequence-verified.
FEATURES
source
1. .1140

[illegible]

from placenta and brain"		
gene	1. .1140	
CDS	/gene="MAPK3"	
	1. .1140	
	/gene="MAPK3"	
	/codon_start=1	
	/transl_table=11	
	/product="mitogen-activated protein kinase 3"	
	/protein_id="AA36307.1"	
	/db_xref="GI:60814578"	
	/translation="MAAAAGGGGGEPRRTGVPGVGEVEMVKGQFPDVGPRTYQ LOVIGBAGTSSAYDHVKRVAIKKISPEHOTYQORTLREIQILLRFRHENVIG IRDILASLTLEARDVYIVQDLMETDLYLKLSQOLSNDCVFLVLOILRGUKYIHS NVLRDLKPSNLLINTTCDLKIKCDPLGARIADPEHDHGFLEFYVATRWYRAPEL SGYTKSIDISVGCILAEMLSNRPFFPGHYLDQNLHILGILGFSFSDLNCIINMK ARNYLOSLESKTKVAWAKLFPKSDSKDLDDRLMTFNPKNKITVEEALAHPLYEQY DPTDEVAEPEPTFAMELDDPLKRLKELIFQETARFQGVLEAP"	
ORIGIN		
Query Match 62.1%; Score 1140; DB 11; Length 1140;		
Best Local Similarity 100.0%; Pred. No. 4.3e-190;		
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	12	ATGGCGGCGCGCGGCTCAGGGGGCGGGGGGAGCCCGTAGAACCCGAGGGGGTC 71
DB	1	ATGGCGGCGCGCGGCTCAGGGGGCGGGGGGAGCCCGTAGAACCCGAGGGGGTC 60
QY	72	GGCCCGGGGGTCCCGGGGAGTGGAGATGGTGAAGGGGCGAGCGGTTGCACTGGGCGG 131
DB	61	GGCCCGGGGGTCCCGGGGAGTGGAGATGGTGAAGGGGCGAGCGGTTGCACTGGGCGG 120
QY	132	CGGTACACGAGTTGCAGTACATCGGCGAGGGGGCGGTACGGCATGGTCACTCGGCTAT 191
DB	121	CGGTACACGAGTTGCAGTACATCGGCGAGGGGGCGGTACGGCATGGTCACTCGGCTAT 180
QY	192	GACCACTGCGCAAGACTCGCTGGCCATCAAGAAGATCAGCCCTTCGAACATCAGACC 251
DB	181	GACCACTGCGCAAGACTCGCTGGCCATCAAGAAGATCAGCCCTTCGAACATCAGACC 240
QY	252	TACTGCGAGCGCAGCTCCGGGAGATCCAGATCTCTGTGGCTTCGGCCATGAGAAATGC 311
DB	241	TACTGCGAGCGCAGCTCCGGGAGATCCAGATCTCTGTGGCTTCGGCCATGAGAAATGC 300
QY	312	ATCGGATCCGAGACATTCGCGGGCGTCCACCTGGAGGCCATGAGATGTCTACATT 371
DB	301	ATCGGATCCGAGACATTCGCGGGCGTCCACCTGGAGGCCATGAGATGTCTACATT 360
QY	372	GTGCAAGACTGTAGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAGCTGAGCAAT 431
DB	361	GTGCAAGACTGTAGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAGCTGAGCAAT 420
QY	432	GACCATATCTGTCTTCTTACAGATCTCTGGGGGCTCAAGTACATCCACTCGGCC 491
DB	421	GACCATATCTGTCTTCTTCTTACAGATCTCTGGGGGCTCAAGTACATCCACTCGGCC 480
QY	492	AACGTGCTCCACCGAGATCTAAAGCCCTCAACCTGCTCATCAACACCCCTGCGACCTT 551
DB	481	AACGTGCTCCACCGAGATCTAAAGCCCTCAACCTGCTCATCAACACCCCTGCGACCTT 540
QY	552	AAGATTGTGATTTCGGCTGGCCGGATTGGCGATCTGAGCATGACACACCGGCTTC 611
DB	541	AAGATTGTGATTTCGGCTGGCCGGATTGGCGATCTGAGCATGACACACCGGCTTC 600
QY	612	CTGACGGAGTATGCGCTACGCGTGTACCGGGCCCGAGATCATGTGAATCCAAAG 671
DB	601	CTGACGGAGTATGCGCTACGCGTGTACCGGGCCCGAGATCATGTGAATCCAAAG 660
QY	672	GGCTATACCAAGTCCATCGACATCTGGTCTGTGGGTGCAATTCGGCTGAGATGCTCT 731
DB	661	GGCTATACCAAGTCCATCGACATCTGGTCTGTGGGTGCAATTCGGCTGAGATGCTCT 720
QY	732	AACGGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAACATTTCTGGGCATC 791

Search completed: February 7, 2006, 15:20:38
Job time : 9916.9 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:22:44 ; Search time 1117.96 Seconds
(without alignments)
10951.209 Million cell updates/sec

Title: US-10-623-108-5

Perfect score: 1837

Sequence: 1 gagagtgagatggcgccg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2010s.*

5: Geneseqn2020s.*

6: Geneseqn2030s.*

7: Geneseqn2040s.*

8: Geneseqn2050s.*

9: Geneseqn2060s.*

10: Geneseqn2070s.*

11: Geneseqn2080s.*

12: Geneseqn2090s.*

13: Geneseqn2100s.*

14: Geneseqn2110s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1837	100.0	1837	14	ADW12904 Human SMA
2	1725	93.9	1813	13	ACF87406 Human SIR
3	1708.4	93.0	1866	2	AAV62478 Human MAP
4	1708.4	93.0	1866	6	ABK84604 Human CDN
5	1708.4	93.0	1866	8	ACA89915 Gene diff
6	1708.4	93.0	1866	12	ADO24424 Human PRO
7	1708.4	93.0	1866	14	ADZ10042 Human bre
8	1708.4	93.0	1866	6	ABK90802 DNA encod
9	1707	92.9	1777	14	ADW12906 Human SMA
10	1694	92.2	1868	4	AAI59696 Human pol
11	1679.8	91.4	1869	4	AAI57910 Human pol
12	1595	86.8	1726	14	ADW12902 Human SMA
13	1451	79.0	1654	14	ADW12900 Human SMA
14	1437	78.2	1699	12	ADH48413 Human KPP
15	1138.4	62.0	1140	10	ADH59631 Erkl enco
16	1138.4	62.0	1896	2	AAV71022 Erkl-gree
17	1138.4	62.0	1896	3	ADAO0036 EGFP-Erkl
18	1128	61.4	1305	13	ACN41199 Human dia
19	1098	59.8	1290	13	ACN41200 Human dia

20	1080.2	58.8	1125	10	ABZ77131 Human pro
21	1076	58.6	1214	3	AAC99095 Human pan
22	1066.8	58.1	1071	2	ABA96212 Extracell
23	1061	57.8	1783	14	ADX26048 Novel cel
24	1043.2	56.8	1147	8	ABX72224 Human NOV
25	1021.8	55.6	1747	2	AAQ20260 ERK1. 8/2
26	1016	55.3	1254	13	ACN41201 Human dia
27	986	53.7	1239	13	ACN41202 Human dia
28	965.4	52.6	1238	13	ADV41142 Rat cardl
29	965.4	52.6	1238	14	ADX26182 Novel cel
30	958.6	52.2	1875	2	AAV71030 Erkl-gree
31	928.4	50.5	1210	13	ACN41203 Human dia
32	651.4	35.5	1484	12	ADQ86026 Human tum
33	649.8	35.4	2934	13	ADP56267 Human PRO
34	649.8	35.4	2934	13	ADP56267 Human PRO
35	649.8	35.4	2934	14	ADY15945 DNA encod
36	646.6	35.2	1611	2	AAV62479 Human MAP
37	646.6	35.2	1611	6	ABL59534 Human mit
38	646.6	35.2	1611	6	ABK90803 DNA encod
39	646.6	35.2	1611	8	ACFI2891 Human cer
40	646.6	35.2	1611	14	ADY20280 DNA encod
41	646	35.2	3409	10	ADF81745 Leukaemia
42	642	34.9	2791	6	ABK84759 Human CDN
43	642	34.9	2791	10	ACA56665 Human sig
44	642	34.9	2791	12	ADI56461 Human pol
45	641.8	34.9	1083	14	ADV42655 Human psy

ALIGNMENTS

RESULT 1

ADW12904

ID ADW12904 standard; cDNA; 1837 BP.

AC ADW12904;

XX 07-APR-2005 (first entry)

XX Human SMAPK3V3 variant gene.

XX DNA purification; diagnosis; cancer; mitogen-activated protein kinase;

XX Gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 12..1151

XX /*tag= a

XX /product= "SMAPK3V3 protein"

XX US2005013817-A1.

XX 20-JAN-2005.

XX 18-JUL-2003; 2003US-00623108.

XX 18-JUL-2003; 2003US-00623108.

XX (DAIK/) DAI K.

XX Dai K;

XX WPI; 2005-080923/09.

XX P-PSDB; ADW12905.

XX New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing

XX diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene

XX in a mammal.

XX Claim 4; SEQ ID NO 5; 55pp; English.

XX The invention relates to novel isolated polypeptide (I) comprising an

CC

QY 1321 GTGAGCATGGCGCTGCCACCTCTCTCTTTCTGAGGCTCCAGCTTCAGGAGGCCA 1380
DB |||||
QY 1381 AGGCTTTCTCTCCCAACCGCCCTCCCAACGGGGCTCGGGAGCTCAGGTGCCCCAGT 1440
DB |||||
QY 1441 TCAATCTCCGCTGCTGCTGCTGCGGCTTACCTTCCCAAGCTCCAGTCTCTGCGAGT 1500
DB |||||
QY 1501 TCTGGAATGAAGGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1560
DB |||||
QY 1561 GCGCTGAGTGAAGTCAAGGCGCATGCTGCGCCCTCATCTCATTTCAAAACCCACCT 1620
DB |||||
QY 1621 AGTTTCCCTGAAGAAATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCG 1680
DB |||||
QY 1681 GCGCGAATCCCTCCCTGTCAAGCTGTCACTTCGCGTGCCTCGCTGCTGCTGCTG 1740
DB |||||
QY 1741 GTGAGCAGAAGTGGAGTGGGGGCGTGGAGAGCCCGGCGCCCTGCCACCTCCCTGACC 1800
DB |||||
QY 1801 CGTCTAATATATAATATAGAGATGTGTCTATGGCTG 1837
DB |||||
QY 1837 CGTCTAATATATAATATAGAGATGTGTCTATGGCTG 1786
DB |||||

RESULT 3
AAV62478
ID AAV62478 standard; cdna; 1866 BP.
XX
AC AAV62478;
XX
DT 18-JAN-1999 (first entry)
XX
DE Human MAP kinase, ERK-1 cDNA sequence.
XX
KW ERK-1; ERK-2; mitogen-activated protein kinase; MAP kinase; human;
KW inhibition; malignant; neoplastic growth; epithelial cell; mammal;
KW endothelial cell; antisense oligonucleotide; primary cancer;
KW metastatic cancer; breast cancer; prostate cancer; angiosarcoma;
KW endocrine tissue cancer; ds.
XX
OS Homo sapiens.
XX
PN W09844101-A1.
XX
PD 08-OCT-1998.
XX
PF 19-MAR-1998; 98WO-US005471.
XX
PR 28-MAR-1997; 97US-00827520.
PR 01-APR-1997; 97US-00831994.
PR 12-AUG-1997; 97US-00909742.
XX
PA (UNYV) UNIV NEW YORK STATE RES FOUND.
XX
PI Sivaraman VS, Wang H, Malbon CC;
XX
DR WPI; 1998-557109/47.
XX
PT Treatment of e.g. breast or prostate cancer or angiosarcoma - by
PT administering antisense oligonucleotides to genes encoding mitogen-

PT activating protein kinases ERK1 and ERK2.
XX
PS Disclosure; Page 11-12; 59pp; English.
XX
CC This represents the cDNA sequence of the human ERK-1, a mitogen-activated
CC protein (MAP) kinase. The invention provides a method of inhibiting
CC malignant neoplastic growth of epithelial or endothelial cell in a mammal
CC which comprises administering to the mammal an effective amount of an
CC oligonucleotide complementary to part of the mRNA for the MAP kinases,
CC ERK-1 or ERK2 which is over-expressed in the mammal. Also provided is a
CC method for identifying and monitoring potentially malignant neoplastic
CC cells by measuring the levels of ERK1 and ERK2 mRNA in epithelial or
CC endothelial cells and comparing it to the levels from normal cells of the
CC same origin. Administration of the ERK1 and ERK2 antisense
CC oligonucleotides (AAV62480 and AAV62481) to neoplastic endothelial or
CC epithelial cells inhibits over-expression of ERK1 and ERK2. This can be
CC used to treat epithelial and endothelial malignancies including primary
CC or metastatic cancers of e.g. the breast, prostate, other endocrine
CC tissue or angiosarcoma
XX
SQ Sequence 1866 BP; 380 A; 605 C; 535 G; 346 T; 0 U; 0 Other;
Query Match 93.0%; Score 1708.4; DB 2; Length 1866;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;
QY 3 GGAGTGGAGATGGCGCGCGCGCTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 62
DB |||||
QY 64 GGAGTGGAGATGGCGCGCGCGCTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 123
DB |||||
QY 63 GAGGGGGTTCGG 122
DB |||||
QY 124 GAGGGGGTTCGG 183
QY 123 GTGG 182
DB |||||
QY 184 GTGG 243
QY 183 TCGGGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAA 242
DB |||||
QY 244 TCGGGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAA 303
QY 243 CATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCCTGCTGCGCTTCGCGCAT 302
DB |||||
QY 304 CATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCCTGCTGCGCTTCGCGCAT 363
QY 303 GAGAAATGTCATCGGCATCCGAGACATTCGCGGGGTCACCCCTGGAAGCCATGAGAGAT 362
DB |||||
QY 364 GAGAAATGTCATCGGCATCCGAGACATTCGCGGGGTCACCCCTGGAAGCCATGAGAGAT 423
QY 363 GTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGAG 422
DB |||||
QY 424 GTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGAG 483
QY 423 CTGAGCAATGACCATATCTGCTACTTCTCTACAGATCCTCGGGGGCTCAAGTACATC 482
DB |||||
QY 484 CTGAGCAATGACCATATCTGCTACTTCTCTACAGATCCTCGGGGGCTCAAGTACATC 543
QY 483 CATCTCGCAACGTGCTCCACGAGATCTAAGCCCTCCAACTGCTCATCAACACCAC 542
DB |||||
QY 544 CATCTCGCAACGTGCTCCACGAGATCTAAGCCCTCCAACTGCTCATCAACACCAC 603
QY 543 TCGGACCTTAAGATTTGTGATTTTCGGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 602
DB |||||
QY 604 TCGGACCTTAAGATTTGTGATTTTCGGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 663
QY 603 ACCGGCTTCTGACGAGTATGTGGCTACGCGCTGTGTACCGGGGGGGGGGGGGGGGGGGGG 662
DB |||||
QY 664 ACCGGCTTCTGACGAGTATGTGGCTACGCGCTGTGTACCGGGGGGGGGGGGGGGGGGGGG 723
QY 663 AACTCAAGGGCTATACCAAGTCCATCGCATCTGTGTGGGGTGCATTCGTGGCTGAG 722
DB |||||
QY 724 AACTCAAGGGCTATACCAAGTCCATCGCATCTGTGTGGGGTGCATTCGTGGCTGAG 783
DB |||||

QY 723 ATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 782
Db 784 ATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 843
QY 783 CTGGGCACTCTGGGCTCCCAATCCAGAGGACCTGAATTGTATCATCAATGAAGGCC 842
Db 844 CTGGGCACTCTGGGCTCCCAATCCAGAGGACCTGAATTGTATCATCAATGAAGGCC 903
QY 843 CGAACTACCTACAGTCTCTGCGCTCCCAAGACCAAGGTGGCTGGGCAAGCTTTTCCCC 902
Db 904 CGAACTACCTACAGTCTCTGCGCTCCCAAGACCAAGGTGGCTGGGCAAGCTTTTCCCC 963
QY 903 AAGTCAGACTCCAAAGCGCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAA 962
Db 964 AAGTCAGACTCCAAAGCGCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAA 1023
QY 963 CGGATCAAGTGGAGAGCGCTGGCTCAACCTTACCTGGAGCACTATGACCCGACG 1022
Db 1024 CGGATCAAGTGGAGAGCGCTGGCTCAACCTTACCTGGAGCACTATGACCCGACG 1083
QY 1023 GATGAGCCAGTGGCGGAGGAGCCTTCACTTCCGATGGAGTGGATGACCTACCTAAG 1082
Db 1084 GATGAGCCAGTGGCGGAGGAGCCTTCACTTCCGATGGAGTGGATGACCTACCTAAG 1143
QY 1083 GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTCTGGAG 1142
Db 1144 GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTCTGGAG 1203
QY 1143 GCGCCCTAGCCAGACAGACATCTCTGCAACCTGGGGCTGGAAACAGAAAGAG 1202
Db 1204 GCGCCCTAGCCAGACAGACATCTCTGCAACCTGGGGCTGGAAACAGAAAGAG 1243
QY 1203 GCAAGAGTCACTGAGGGCTCTGTCAACCAAGGAGCTGCTCTGCTGCGCTCTCCCG 1262
Db 1244 -----GGACCTGCTCTCTGCTGCGCTCTCTCCCG 1272
QY 1263 CCAGACTGTTAGAAAATGACACTGTGCGCCAGCCGAGCCTTGGCAGCCAGGCGGGGT 1322
Db 1273 CCAGACTGTTAGAAAATGACACTGTGCGCCAGCCGAGCCTTGGCAGCCAGGCGGGGT 1332
QY 1323 GGAGATGGGCTGGCCCACTCTCTCTCTTTGCTGAGGCTTCAGCTTCAGGAGCCAG 1382
Db 1333 GGAGATGGGCTGGCCCACTCTCTCTCTTTGCTGAGGCTTCAGCTTCAGGAGCCAG 1392
QY 1383 GCCTTCTCTCCCAACCGCCCTCCCAACCGGGCTCGGAGCTCAGTGGCCCAAGTTC 1442
Db 1393 GCCTTCTCTCCCAACCGCCCTCCCAACCGGGCTCGGAGCTCAGTGGCCCAAGTTC 1452
QY 1443 AATCTCCC ---GCTGCTGCTGCTGCGCCCTTACCTTCCAGCGCTCCAGTCTCTGGCAG 1499
Db 1453 AATCTCCGCTGCTGCTGCTGCTGCGCCCTTACCTTCCAGCGCTCCAGTCTCTGGCAG 1512
QY 1500 TTCTGGAATGAAGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGTGGAGGGTGG 1559
Db 1513 TTCTGGAATGAAGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGTGGAGGGTGG 1572
QY 1560 GGGCGCTAGTAGGACTCAGGGCCTGCTGCGCCCTCTCATCTCATCAACCCCAACC 1619
Db 1573 GGGCGCTAGTAGGACTCAGGGCCTGCTGCGCCCTCTCATCTCATCAACCCCAACC 1632
QY 1620 TAGTTTCTCTGAAGGAATCTTCTAGTCTCAAGGGCTAGCATCCTGAGGAGCCAGGCC 1679
Db 1633 TAGTTTCTCTGAAGGAATCTTCTAGTCTCAAGGGCTAGCATCCTGAGGAGCCAGGCC 1692
QY 1680 GGGCGGAATCCCTCTCTCAAGCTGTCACTTCGCTGCGCTGCTGCTCTCTGTGTGT 1739
Db 1693 GGGCGGAATCCCTCTCTCAAGCTGTCACTTCGCTGCGCTGCTGCTCTCTGTGTGT 1752
QY 1740 GGTGAGCAAGTGAAGCTGGGGGCTGAGAGCCCGCGCCCTGCGCACTCCCTGAC 1799
Db 1753 GGTGAGCAAGTGAAGCTGGGGGCTGAGAGCCCGCGCGCTGCGCACTCCCTGAC 1812

QY 1800 CCCTTAATATATAATATAGAGATGTCTATGGCTG 1837
Db 1813 CCCTTAATATATAATATAGAGATGTCTATGGCTG 1850

RESULT 4

ABK84604
ID ABK84604 standard; cDNA; 1866 BP.
AC ABK84604;
XX
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #1175.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.

OS Homo sapiens.

XX WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US030821.

XX 03-OCT-2000; 2000US-0237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 1175; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's

CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1866 BP; 380 A; 605 C; 535 G; 346 T; 0 U; 0 Other;

Query Match 93.0%; Score 1708.4; DB 6; Length 1866;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;

QY 3 GGAGTGGAGATGCGCGCGCGCGCGGTCTAGGGGGGCGGGGCGGGAGCCCGTAGAAC 62
DB 64 GGAGTGGAGATGCGCGCGCGCGCGGTCTAGGGGGGCGGGGCGGGAGCCCGTAGAAC 123

QY 63 GAGGGGTGCGCGCGCGCGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTGAC 122
DB 124 GAGGGGTGCGCGCGCGCGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTGAC 183

QY 123 GTGGGCGCGGCTACACGCGAGTTGCGAGTACATCGGGAGGGCGGTACGGCATGGTTCAGC 182
DB 184 GTGGGCGCGGCTACACGCGAGTTGCGAGTACATCGGGAGGGCGGTACGGCATGGTTCAGC 243

QY 183 TCGGGCTATGACACGCTGCGCAAGACTCGCGTGGCCATCAAGAGATACGCCCCCTTCGAA 242
DB 244 TCGGCTATGACACGCTGCGCAAGACTCGCGTGGCCATCAAGAGATACGCCCCCTTCGAA 303

QY 243 CATCAGACCTACTCCGAGCGACGCTCCGGAGATCCAGATCTGCTGCGCTTCGCGCAT 302
DB 304 CATCAGACCTACTCCGAGCGACGCTCCGGAGATCCAGATCTGCTGCGCTTCGCGCAT 363

QY 303 GAGATGTCATTCGGCATCCCGAGACATTCGCGGCGTCCACCTCGAAGCCATGAGAGAT 362
DB 364 GAGATGTCATTCGGCATCCCGAGACATTCGCGGCGTCCACCTCGAAGCCATGAGAGAT 423

QY 363 GTCTAATTGTGAGGACCTGATGGAGACTGACCTGTACAAAGTTGCTGAAAGCCAGCAG 422
DB 424 GTCTAATTGTGAGGACCTGATGGAGACTGACCTGTACAAAGTTGCTGAAAGCCAGCAG 483

QY 423 CTGAGCAATGACCATATCTGCTACTTCTCTACAGATCTGCGGGGCTCAAGTACATC 482
DB 484 CTGAGCAATGACCATATCTGCTACTTCTCTACAGATCTGCGGGGCTCAAGTACATC 543

QY 483 CACTCCGCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACC 542
DB 544 CACTCCGCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCAGCAACACCACC 603

QY 543 TCGGACCTTAAGATTTGTGATTTTCGGCTTGGCCGCGGATTCGCCATCTGAGCATGACAC 602
DB 604 TCGGACCTTAAGATTTGTGATTTTCGGCTTGGCCGCGGATTCGCCATCTGAGCATGACAC 663

QY 603 ACCGGCTTCCTGACGAGTATGGCTTACCGGTGTTACCGGGCCCGCAGAGATCATGCTG 662
DB 664 ACCGGCTTCCTGACGAGTATGGCTTACCGGTGTTACCGGGCCCGCAGAGATCATGCTG 723

QY 663 AACTCCAAAGGGCTATACCAAGTCCATCGCATCTGCTGCTGCGGCTGCAATTCGGCTGAG 722
DB 724 AACTCCAAAGGGCTATACCAAGTCCATCGCATCTGCTGCTGCGGCTGCAATTCGGCTGAG 783

QY 723 ATGCTCTCTAACCGGGCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCAATT 782
DB 784 ATGCTCTCTAACCGGGCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCAATT 843

QY 783 CTGGGCATCTGGGCTCCCATCCGAGGAGCTGAATTTGTATCATCAACCAAGGCC 842
DB 844 CTGGGCATCTGGGCTCCCATCCGAGGAGCTGAATTTGTATCATCAACCAAGGCC 903

QY 843 CGAAACTACATGCTCTGCGCTCCCAAGCAAGGTGGCTTTGGGCGAAGCTTTTCCCG 902

DB 904 CGAAACTACCTACAGTCTCTGCGCTCCCAAGACCAAGTGGCTTGGGCCCAAGCTTTTCCCG 963
QY 903 AAGTCAGATCTCCAAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAA 962
DB 964 AAGTCAGATCTCCAAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAA 1023
QY 963 CGGATCACAGTGGAGGAAGCGTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGAGC 1022
DB 1024 CGGATCACAGTGGAGGAAGCGTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGAGC 1083

QY 1023 GATGAGCAGTGGCCGAGAGGCCCTTACCTTCGCCATGGAGCTGGATGACCTACTACTAAG 1082
DB 1084 GATGAGCAGTGGCCGAGAGGCCCTTACCTTCGCCATGGAGCTGGATGACCTACTACTAAG 1143

QY 1083 GAGCGGCTGAAGAGAGCTCATCTTCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1142
DB 1144 GAGCGGCTGAAGAGAGCTCATCTTCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1203

QY 1143 GCGCCCTAGCCCGACAGACATCTCTGCAACCTCGGGGCTCGGAACAGAACTGGCAAGAG 1202
DB 1204 GCGCCCTAGCCCGACAGACATCTCTGCAACCTCGGGGCT----- 1243

QY 1203 GCAAGAGTCACTGAGGGGCTCTGTCAACCCAGGACCTGCTCTGCTGCTCCCTCTCCCG 1262
DB 1244 -----GGACCTGCTCTGCTGCTCTCCCG 1272

QY 1263 CCAGACTGTTAGAAATGGACACTGTGCCAGCCGCGACCTTGGCAGCCCGCGGGGT 1322
DB 1273 CCAGACTGTTAGAAATGGACACTGTGCCAGCCGCGACCTTGGCAGCCCGCGGGT 1332

QY 1323 GGAGCATGGGCTGGCCACCTCTCTTCTTGTGAGGGCTCCAGCTTTCAGGCAGGCGCAAG 1382
DB 1333 GGAGCATGGGCTGGCCACCTCTCTTCTTGTGAGGGCTCCAGCTTTCAGGCAGGCGCAAG 1392

QY 1383 GCGTTCTCTCCCGACCGGCTTCCCGACGCGGGCTTCGGAAGCTCAGGTGGGCGCCAGTTC 1442
DB 1393 GCGTTCTCTCCCGACCGGCTTCCCGACGCGGGCTTCGGAAGCTCAGGTGGGCGCCAGTTC 1452

QY 1443 AATCTCCC---GCTGCTGCTGCTGCGGCTTACCTTCCCGACGCTCCAGTCTCTGGCAG 1499
DB 1453 AATCTCCC---GCTGCTGCTGCTGCGGCTTACCTTCCCGACGCTCCAGTCTCTGGCAG 1512

QY 1500 TTCTGGAATGGAAGGGTCTGGCTGCGCCCAACCTTCTGTAAGGGGCGAGGTGGAGGGTGG 1559
DB 1513 TTCTGGAATGGAAGGGTCTGGCTGCGCCCAACCTTCTGTAAGGGGCGAGGTGGAGGGTGG 1572

QY 1560 GGGCGCTGATGAGGACTCAGGGCCATGCTGCGCCCTCATCTCATTTCAAACCCCAACC 1619
DB 1573 GGGCGCTGATGAGGACTCAGGGCCATGCTGCGCCCTCATCTCATTTCAAACCCCAACC 1632

QY 1620 TAGTTTCCCTGGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCC 1679
DB 1633 TAGTTTCCCTGGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCC 1692

QY 1680 GGGCGGAATCCCTCTGCTCAAAGCTGTCACTTTCGCGTGGCTCGCTCTCTGTGTGT 1739
DB 1693 GGGCGGAATCCCTCTGCTCAAAGCTGTCACTTTCGCGTGGCTCGCTCTCTGTGTGT 1752

QY 1740 GGTGAGCAAAATGAGCTGGGGGGGTGGAGAGCCCGGCGCCCTTCCCACTCCCTGAC 1799
DB 1753 GGTGAGCAAAATGAGCTGGGGGGGTGGAGAGCCCGGCGCCCTTCCCACTCCCTGAC 1812

QY 1800 CCGTCTTAATATAAATATAGAGATGTGTCTATGCGTG 1837
DB 1813 CCGTCTTAATATAAATATAGAGATGTGTCTATGCGTG 1850

RESULT 5
AC89915
ID AC89915 standard; cDNA; 1866 BP.
XX AC
XX AC89915;

Db	544	CACTCGGCAACGTCCTCCACCGAGATCTAAAGCCCTCAACCTGCTCAACAACCAACC	603
Qy	543	TGCGACCTTTAGATTGTGATTTGGCGCTGGCCCGGATTCGCCATCTCTGAGCATGACCAC	602
Db	604	TGCGACCTTTAGATTGTGATTTGGCGCTGGCCCGGATTCGCCATCTCTGAGCATGACCAC	663
Qy	603	ACCGGCTTCTTGACCGGAGTATGTGGCTACGCGCTGGTACCGGGGCCCAGAGATCATGCTG	662
Db	664	ACCGGCTTCTTGACCGGAGTATGTGGCTACGCGCTGGTACCGGGGCCCAGAGATCATGCTG	723
Qy	663	AACTCAAAGGGCTATAACAAGTCCATCGACAATCTGTGCTCTGTGGCTGCATTTCTGGCTGAG	722
Db	724	AACTCAAAGGGCTATAACAAGTCCATCGACAATCTGTGCTCTGTGGCTGCATTTCTGGCTGAG	783
Qy	723	ATGCTCTCTTAACCGGCCCATCTTCCCTGCGAAGCACTACCTGGATCAGCTCAACCAATT	782
Db	784	ATGCTCTCTTAACCGGCCCATCTTCCCTGCGAAGCACTACCTGGATCAGCTCAACCAATT	843
Qy	783	CTGGGCATCTGGGCTCCCCATCCAGGAGGACCTGAAATTGTATCATCAACATGAAGGCC	842
Db	844	CTGGGCATCTGGGCTCCCCATCCAGGAGGACCTGAAATTGTATCATCAACATGAAGGCC	903
Qy	843	CGAAACTTACTTACGTCTCTGCTTCCGCTTCAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC	902
Db	904	CGAAACTTACTTACGTCTCTGCTTCCGCTTCAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC	963
Qy	903	AGTTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAA	962
Db	964	AGTTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAA	1023
Qy	963	CGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG	1022
Db	1024	CGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG	1083
Qy	1023	GATGAGCCATGTGGCCGAGAGCCCTTCACTTGGCCATPFGAGCTTGGATGACTTACCTAAG	1082
Db	1084	GATGAGCCATGTGGCCGAGAGCCCTTCACTTGGCCATPFGAGCTTGGATGACTTACCTAAG	1143
Qy	1083	GAGCGGCTGAAAGGACTCATCTTCCAGAGACAGCAGCGCTTCCAGCCCGGAGTGTCTGGAG	1142
Db	1144	GAGCGGCTGAAAGGACTCATCTTCCAGAGACAGCAGCGCTTCCAGCCCGGAGTGTCTGGAG	1203
Qy	1143	GGCCCTAGCCACAGACAGACATCTCTGCAACCTGGGGCTGGAAACAGAACTGGCAAGAG	1202
Db	1204	GGCCCTAGCCACAGACAGACATCTCTGCAACCTGGGGCTGGAAACAGAACTGGCAAGAG	1243
Qy	1203	GCAAGAGTCACTGAGGGCTCTGTCAACCCAGGACCTTGCCTTCTGCTTGCCCTCTCCCG	1262
Db	1244	-----GGACCTGCTTCTGCTTGCCCTCTCCCG	1272
Qy	1263	CCGACTGTTAGAAATGGACACTGTGCCAGCCGCGACCTTGGCAGCCCGAGCGCGGGT	1322
Db	1273	CCGACTGTTAGAAATGGACACTGTGCCAGCCGCGACCTTGGCAGCCCGAGCGCGGGT	1332
Qy	1323	GGAGCATGGGCTGGGCACTCTCTCTTTGCTGAGGCTCCAGCTTCAGGACAGCCCAAG	1382
Db	1333	GGAGCATGGGCTGGGCACTCTCTCTTTGCTGAGGCTCCAGCTTCAGGACAGCCCAAG	1392
Qy	1383	GCCTTCTCTCCCAACCCGCCCTCCCAAGGGGCTCTGGGAGCTCAGGTGGCCCAAGTTC	1442
Db	1393	GCCTTCTCTCCCAACCCGCCCTCCCAAGGGGCTCTGGGAGCTCAGGTGGCCCAAGTTC	1452
Qy	1443	AATCTCCC--GCTGCTGCTGCGCCCTTACCTTCCAGCGTCCAGCTCTCTGCGAG	1499
Db	1453	AATCTCCCCTGCTGCTGCTGCGCCCTTACCTTCCAGCGTCCAGCTCTCTGCGAG	1512
Qy	1500	TTCTGGAATGGAAGGTTCTGGCTGCCCAACTCTGCTGAAAGGGCAGAGGTTGGAGGTTGGG	1559
Db	1513	TTCTGGAATGGAAGGTTCTGGCTGCCCAACTCTGCTGAAAGGGCAGAGGTTGGAGGTTGGG	1572
Qy	1560	GGGCGCTGAGTAGGGAATCAGGGCCATGCTTCCGCCCTCTCATCTCATTAACCCCAACC	1619

Query Match	93.0%	Score 1708.4;	DB 14;	Length 1866;
Best Local Similarity	97.0%;	Pred. No. 0;		
Matches 1783;	Conservative	0;	Mismatches	1;
Indels	54;	Gaps	2;	

QY 3 GGAGTGGAGATGGCGCGCGCGCGGCTCAGGGGGGCGGGGCGGGAGCGCCCGTAGAAC 62
DB 64 GGAGTGGAGATGGCGCGCGCGCGGCTCAGGGGGGCGGGGCGGGAGCGCCCGTAGAAC 123
QY 63 GAGGGGGTCCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTTCGAC 122
DB 124 GAGGGGGTCCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTTCGAC 183
QY 123 GTGGGCGCGCTACACGAGTTTGAGTACATCGGCGAGGGCGGCTACGGCATGTTGTCAG 182
DB 184 GTGGGCGCGCTACACGAGTTTGAGTACATCGGCGAGGGCGGCTACGGCATGTTGTCAG 243
QY 183 TCGGCTATGACCACTGTGGCAGACTCGGGTGGCCATCAAGAGATCAGCCCTTCGAA 242
DB 244 TCGGCTATGACCACTGTGGCAGACTCGGGTGGCCATCAAGAGATCAGCCCTTCGAA 303
QY 243 CATCAGACCTACTCGCCAGGGCAGCTCCGGGAGATCCAGATCCTGTGGGCTTCGCCCAT 302
DB 304 CATCAGACCTACTCGCCAGGGCAGCTCCGGGAGATCCAGATCCTGTGGGCTTCGCCCAT 363
QY 303 GAGAAATGTCATCGGCATCCGAGACATTCGTGGGGCGTCCACCTCGAAGCCATGAGAGAT 362
DB 364 GAGAAATGTCATCGGCATCCGAGACATTCGTGGGGCGTCCACCTCGAAGCCATGAGAGAT 423
QY 363 GTCTACATTTGTGAGGACCTGATGGAGACTGAGCTGTACAGTTGCTGAAGCCAGCAG 422
DB 424 GTCTACATTTGTGAGGACCTGATGGAGACTGAGCTGTACAGTTGCTGAAGCCAGCAG 483
QY 423 CTGAGCAATGACCATATCTGCTACTTTCCTACAGATCTCGGGGGCTCAAGTACATC 482
DB 484 CTGAGCAATGACCATATCTGCTACTTTCCTACAGATCTCGGGGGCTCAAGTACATC 543
QY 483 CACTCCGCCAACGTGTCTCAACGAGATCTAAGGCCCTCCAACTGTCTATCAACACCAC 542
DB 544 CACTCCGCCAACGTGTCTCAACGAGATCTAAGGCCCTCCAACTGTCTATCAACACCAC 603
QY 543 TGGGACCTTAAGATTTGTGATTTGGCGCTGGCGCGGATTTGCCGATCTGAGCATGACAC 602
DB 604 TGGGACCTTAAGATTTGTGATTTGGCGCTGGCGCGGATTTGCCGATCTGAGCATGACAC 663
QY 603 ACCGGCTCTCTGACGAGTATGTGGCTACGCGCTGTGTACCGGGCCCGCAGAGATCATGCTG 662
DB 664 ACCGGCTCTCTGACGAGTATGTGGCTACGCGCTGTGTACCGGGCCCGCAGAGATCATGCTG 723
QY 663 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGTGATTTCTGGCTGAG 722
DB 724 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGTGATTTCTGGCTGAG 783
QY 723 ATGCTCTCTAAACGGGCCATCTTCCCTGGCAGCAGCTACCTGGATCAGCTCAACACCAT 782
DB 784 ATGCTCTCTAAACGGGCCATCTTCCCTGGCAGCAGCTACCTGGATCAGCTCAACACCAT 843
QY 783 CTGGGATCTCTGGGCTCCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAAGGCC 842
DB 844 CTGGGATCTCTGGGCTCCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAAGGCC 903
QY 843 CGAAACTACTACAGTCTCTGCGCTCCAAAGACCAAGGGGCTTGGGCCAAGCTTTTCCCG 902
DB 904 CGAAACTACTACAGTCTCTGCGCTCCAAAGACCAAGGGGCTTGGGCCAAGCTTTTCCCG 963
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAAACCCCATAA 962
DB 964 AAGTCAGACTCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAAACCCCATAA 1023
QY 963 CGGATCAGTGGAGGAAGCGCTGGGCTCACCCCTACCTGGAGCAGTACTATGACCCGAG 1022
DB 1024 CGGATCAGTGGAGGAAGCGCTGGGCTCACCCCTACCTGGAGCAGTACTATGACCCGAG 1083
QY 1023 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGAGGTGATGACCTTACCTAAG 1082
DB 1084 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGAGGTGATGACCTTACCTAAG 1143

QY 1083 GAGCGGCTGAAGGAGCTCATCTTCAGGAGACAGCACGCTTCAGCGCGGAGTGTCTGGAG 1142
DB 1144 GAGCGGCTGAAGGAGCTCATCTTCAGGAGACAGCACGCTTCAGCGCGGAGTGTCTGGAG 1203
QY 1143 GCCCCTTAGCCAGACAGACATCTCTGTGCACTCTGGGCGCTGGAA CAGAACTGGCAAGAG 1202
DB 1204 GCCCCTTAGCCAGACAGACATCTCTGTGCACTCTGTGCACTCTGGGCGCT----- 1243
QY 1203 GCAAGAGTCACTGAGGGGCTCTGTCA CCGAGGACTGCTCTGCTGCTGCCCTCTCCCG 1262
DB 1244 -----GGA CCTGCTCTCTGCTGCTCTCTCTCCG 1272
QY 1263 CCAGACTGTGAGAAATGACACTGTGCGCCAGCCCGGACCTTGGCAGCCAGCGCGGGT 1322
DB 1273 CCAGACTGTGAGAAATGACACTGTGCGCCAGCCCGGACCTTGGCAGCCAGCGCGGGT 1332
QY 1323 GGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTTCCAGCTTCAGGAGGCGCAAG 1382
DB 1333 GGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTTCCAGCTTCAGGAGGCGCAAG 1392
QY 1383 GCCTTCTCTCCACCGCCCTCCCAAGGGGCTCGGAGCTCAGGTGGGCCCGCAGTTTC 1442
DB 1393 GCCTTCTCTCCACCGCCCTCCCAAGGGGCTCGGAGCTCAGGTGGGCCCGCAGTTTC 1452
QY 1443 AATCTCCC--GCTGCTGCTGCGCCCTTACCTTCCCGAGCGTCCCAGTCTCTGGCAG 1499
DB 1453 AATCTCCCCTGCTGCTGCTGCGCCCTTACCTTCCCGAGCGTCCCAGTCTCTGGCAG 1512
QY 1500 TTCTGGAATGGAAGGGTTCTGGCTGCCCAACCTCTGTAAGGGCAGAGGTGAGGGTGGG 1559
DB 1513 TTCTGGAATGGAAGGGTTCTGGCTGCCCAACCTCTGTAAGGGCAGAGGTGAGGGTGGG 1572
QY 1560 GGGCGCTGATGAGGACTCAGGGCCATGCTGCCCGCTCATCTCAATCAAAACCCACCC 1619
DB 1573 GGGCGCTGATGAGGACTCAGGGCCATGCTGCCCGCTCATCTCAATCAAAACCCACCC 1632
QY 1620 TAGTTTCCCTGAAGAAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1679
DB 1633 TAGTTTCCCTGAAGAAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1692
QY 1680 GGGCCGAATCCCTCCCTGTCAAGCTGTCATTCGCGTGGCCCTCGCTGCTTCTGTGTGT 1739
DB 1693 GGGCCGAATCCCTCCCTGTCAAGCTGTCATTCGCGTGGCCCTCGCTGCTTCTGTGTGT 1752
QY 1740 GGTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCCGCGCCCTGCCACCTCCCTGAC 1799
DB 1753 GGTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCCGCGCCCTGCCACCTCCCTGAC 1812
QY 1800 CCGTCTAATATATAAATATAGAGATGTGTATGGCTG 1837
DB 1813 CCGTCTAATATATAAATATAGAGATGTGTATGGCTG 1850

RESULT 8
ABK90802
ID ABK90802 standard; DNA; 3732 BP.
XX AC ABK90802;
XX DT 05-NOV-2002 (first entry)
XX DE DNA encoding human extracellular signal regulated kinase, ERK1.
XX KW Extracellular signal regulated kinase; hyperalgesia; surgery;
KW opiod withdrawal; pain sensitisation; analgesic; chronic pain; ERK1;
XX human; gene; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 73..1212
FT FT /*tag= a
FT FT /product= "ERK1"

FT	W0200258687-A2.	/note= "Extracellular signal regulated kinase"	
XX	PN		
XX	XX		
XX	PD	01-AUG-2002.	
XX	PF	25-JAN-2002; 2002WO-US0002128.	
XX	PR	25-JAN-2001; 2001US-0264336P.	
XX	PA	(TEXA) UNIV TEXAS SYSTEM.	
XX	PI	Gutstein HB;	
XX	PI	WPI; 2002-608414/65.	
XX	DR	P-PSDB; ABG31847.	
XX	DR		
XX	PT	Reducing or reversing tolerance, physical dependence, hyperalgesia,	
XX	PT	withdrawal symptoms, or pain sensitization in patients on analgesics for	
XX	PT	chronic pain, comprises inhibition of the extracellular signal-regulated	
XX	PT	kinase (ERK).	
XX	PS	Claim 15; Page 138-141; 163pp; English.	
XX	CC	The invention relates to a method of reducing or reversing tolerance,	
XX	CC	reducing the risk of physical dependence or hyperalgesia, reducing the	
XX	CC	symptoms of opioid withdrawal or inhibiting pain sensitisation in a	
XX	CC	patient taking analgesics. The method comprises administering an	
XX	CC	analgesic and an extracellular signal-regulated kinase (ERK) inhibitor	
XX	CC	comprised in a formulation to reduce or reverse tolerance, risk of	
XX	CC	physical dependence, hyperalgesia, symptoms of opioid withdrawal, or	
XX	CC	inhibiting pain sensitisation in patients taking analgesics for chronic	
XX	CC	pain or those undergoing surgery. The present sequence represents the	
XX	CC	coding sequence of human ERK1	
XX	SQ	Sequence 3732 BP; 760 A; 1210 C; 1070 G; 692 T; 0 U; 0 Other;	
		Query Match 93.08; Score 1708.4; DB 6; Length 3732;	
		Best Local Similarity 97.08; Pred. No. 0;	
		Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;	
Qy	3	GGAGTGGAGATGGCGGCGCGCGGCTCAGGGGGCGGGGGCGGGGAGCCCGTAGAAC 62	
Db	64	GGAGTGGAGATGGCGGCGCGCGGCTCAGGGGGCGGGGGCGGGGAGCCCGTAGAAC 123	
Qy	63	GAGGGGTTCGGCCCGGGGTTCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTCGAC 122	
Db	124	GAGGGGTTCGGCCCGGGGTTCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTCGAC 183	
Qy	123	GTGGGCGCGGCTACAGCAGTTCAGTACATCGGCGAGGCGCGGTACGGCATGGTTCAGC 182	
Db	184	GTGGGCGCGGCTACAGCAGTTCAGTACATCGGCGAGGCGGTACGGCATGGTTCAGC 243	
Qy	183	TCGGCCTATGACCAAGTTCGCAAGCTCGGTGGCCATCAAGAAGATCAGCCCTTCGAA 242	
Db	244	TCGGCCTATGACCAAGTTCGCAAGCTCGGTGGCCATCAAGAAGATCAGCCCTTCGAA 303	
Qy	243	CATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCTCTCGGTTCGGCCAT 302	
Db	304	CATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCTCTCGGTTCGGCCAT 363	
Qy	303	GAGAAATGTCATCGGCATCCGAGACATTCGGGGCTCCACCTGGAGGCCATCAGAGAT 362	
Db	364	GAGAAATGTCATCGGCATCCGAGACATTCGGGGCTCCACCTGGAGGCCATCAGAGAT 423	
Qy	363	GTCTACATTTGTCAGGACCTGATGAGACTGACCTGTGTAAGTTGCTGAAAGCCAGAG 422	
Db	424	GTCTACATTTGTCAGGACCTGATGAGACTGACCTGTGTAAGTTGCTGAAAGCCAGAG 483	
Qy	423	CTGAGCAATGACCAATATCTGCTACTTCTTACAGATCTTCGGGGGCTCAAGTACATC 482	
Db	484	CTGAGCAATGACCAATATCTGCTACTTCTTACAGATCTTCGGGGGCTCAAGTACATC 543	
Qy	483	CACTCCGCCAAGCTGTCTCCACCGAGATCTAAAGCCCTCCAACTGTCTATCAACACCAC 542	
Db	544	CACTCCGCCAAGCTGTCTCCACCGAGATCTAAAGCCCTCCAACTGTCTATCAACACCAC 603	
Qy	543	TGGGACCTTAAGATTTGTGATTTCCGCCCTGGCCCGGATTCGGCATCTGAGCATGACCAC 602	
Db	604	TGGGACCTTAAGATTTGTGATTTCCGCCCTGGCCCGGATTCGGCATCTGAGCATGACCAC 663	
Qy	603	ACCGGCTTCTGACCGGAGTATGTGGCTACGCGCTGGTACCGGCCCCAGAGATCATGCTG 662	
Db	664	ACCGGCTTCTGACCGGAGTATGTGGCTACGCGCTGGTACCGGCCCCAGAGATCATGCTG 723	
Qy	663	AACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCATCTTGGGCTGAG 722	
Db	724	AACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCATCTTGGGCTGAG 783	
Qy	723	ATGCTCTTAACCGGCCCATCTTCCCTGGAGCACTACTGATCAGCTCAACACCATTT 782	
Db	784	ATGCTCTTAACCGGCCCATCTTCCCTGGAGCACTACTGATCAGCTCAACACCATTT 843	
Qy	783	CTGGGCTCTTGGGCTTCCCATCCAGAGGAGCTTGAATTTGATATCATCAATGAAGGCC 842	
Db	844	CTGGGCTCTTGGGCTTCCCATCCAGAGGAGCTTGAATTTGATATCATCAATGAAGGCC 903	
Qy	843	CGAACTACTCAGTCTCTGCGCTCCAAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC 902	
Db	904	CGAACTACTCAGTCTCTGCGCTCCAAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC 963	
Qy	903	AAGTCAGCTCCAAAGCCCTTGACCTGGAGCGGATTTAACTTTAAACCCCAATAAA 962	
Db	964	AAGTCAGCTCCAAAGCCCTTGACCTGGAGCGGATTTAACTTTAAACCCCAATAAA 1023	
Qy	963	CGATCACAAGTGGAGGAGCGCTGCTCACCCCTACTCTGGAGCAGTACTATGACCCGACG 1022	
Db	1024	CGATCACAAGTGGAGGAGCGCTGCTCACCCCTACTCTGGAGCAGTACTATGACCCGACG 1083	
Qy	1023	GATGAGCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGATGATGATACCTAAG 1082	
Db	1084	GATGAGCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGATGATGATACCTAAG 1143	
Qy	1083	GAGCGCTGAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1142	
Db	1144	GAGCGCTGAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1203	
Qy	1143	GCCCCCTTAGCCACAGACAGATCTCTGCACCTCTGGGCGCTTGGAAACAGAAAGAG 1202	
Db	1204	GCCCCCTTAGCCACAGACAGATCTCTGCACCTCTGGGCGCTTGGAAACAGAAAGAG 1243	
Qy	1203	GCAAGAGGTCACTGAGGGCCCTCTGTCAACCGAGACCTGCTCTGCTGCGCCCTCTCCCG 1262	
Db	1244	-----GGACCTGCTCTGCTGCTGCTCTCTCCCG 1272	
Qy	1263	CCAGACTGTTAGAAATGGACACTGTGCCAGCCCGGACCTTGGCAGCCCGAGCGGGGT 1322	
Db	1273	CCAGACTGTTAGAAATGGACACTGTGCCAGCCCGGACCTTGGCAGCCCGAGCGGGGT 1332	
Qy	1323	GGAGCATGGGCTTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTTCAGGCGGCAAG 1382	
Db	1333	GGAGCATGGGCTTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTTCAGGCGGCAAG 1392	
Qy	1383	GCCTTCTCTCCCAACCGGCCCTCCCAACCGGGCTTCGGGAGTTCAGGTGGGCCAGTTC 1442	
Db	1393	GCCTTCTCTCCCAACCGGCCCTCCCAACCGGGCTTCGGGAGTTCAGGTGGGCCAGTTC 1452	
Qy	1443	AATCTCCC---GCTGTGCTGCTGGCGCTTACCTTCCCGAGGCTCCAGTCTCTGCGAG 1499	
Db	1453	AATCTCCC---GCTGTGCTGCTGGCGCTTACCTTCCCGAGGCTCCAGTCTCTGCGAG 1512	
Qy	1500	TTCTGGAATGGAAGGTTCTGGGCTGCCCAACCTGCTGAAGGGGAGAGGTGGAGGGTGGG 1559	
Db	1513	TTCTGGAATGGAAGGTTCTGGGCTGCCCAACCTGCTGAAGGGGAGAGGTGGAGGGTGGG 1572	
Qy	1560	GGGCGCTGAGTAGGGACTCAGGGGCAATGCTGCGCCCTCATCTCATTTCAAAACCCACCC 1619	

Db	1573	GGGCGCTGAGTAGGGAGCTCAAGGCGCATGCTGCCCCCTCATCTCATTTCAACCCCAACC	1632
Qy	1620	TAGTTTCCCTGAAGGAACATTCCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC	1679
Db	1633	TAGTTTCCCTGAAGGAACATTCCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC	1692
Qy	1680	GGGCGCGAATCCCTCCCTGTCCTCAAGAGCTGTCACTTCGCGTCCCTCGCTGCTTCTGTGTGT	1739
Db	1693	GGGCGCGAATCCCTCCCTGTCCTCAAGAGCTGTCACTTCGCGTCCCTCGCTGCTTCTGTGTGT	1752
Qy	1740	GTTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGGCGCCCTGCCACCTCCCTGAC	1799
Db	1753	GTTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGGCGCCCTGCCACCTCCCTGAC	1812
Qy	1800	CCGCTTAATATATAAATATAGAGATGTCTATGGCTG	1837
Db	1813	CCGCTTAATATATAAATATAGAGATGTCTATGGCTG	1850
RESULT 9			
ADW12906			
ID	ADW12906 standard; cDNA; 1777 BP.		
XX	ADW12906;		
XX	07-APR-2005 (first entry)		
DT	Human SMAPK3V4 variant gene.		
DE	DNA purification; diagnosis; cancer; mitogen-activated protein kinase;		
XX	gene; ss.		
KW	Homo sapiens.		
KW	Homo sapiens.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Homo sapiens.		
FN	US2005013817-A1.		
XX	20-JAN-2005.		
XX	18-JUL-2003; 2003US-00623108.		
PF	18-JUL-2003; 2003US-00623108.		
XX	(DAIK/) DAI K.		
XX	Dai K;		
XX	WPI; 2005-080923/09.		
DR	P-PSDB; ADW12907.		
XX	New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing		
PT	diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene		
PT	in a mammal.		
XX	Claim 4; SEQ ID NO 7; 55pp; English.		
XX	The invention relates to novel isolated polypeptide (I) comprising an		
CC	amino acid sequence selected from sequences comprising 335 or 359 amino		
CC	acids (ADW12901 or ADW12903) or its fragments. The polypeptides and		
CC	polynucleotides are useful for diagnosing diseases, e.g. cancers,		
CC	associated with the deficiency of the SMAPK3 gene in a mammal. The		
CC	fragments of the polypeptides and polynucleotides can also be used as		
CC	primers or probes. This sequence corresponds to the cDNA encoding the		
CC	SMAPK3V1 variant protein.		
XX	Sequence 1777 BP; 355 A; 575 C; 507 G; 340 T; 0 U; 0 Other;		
XX	Query Match		
XX	92.9%; Score 1707; DB 14; Length 1777;		

Best Local Similarity 96.7%; Pred: No. 0; Mismatches						
Matches 1777; Conservative 0; Indels 60; Gaps 1;						
Qy	1	GAGGAGTGGAGATGCGGGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGACCCCGGTAGAA	60			
Dd	1	GAGGAGTGGAGATGCGGGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGACCCCGGTAGAA	60			
Qy	61	CCGAGGGGTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGTGTGAAGGGGCAGCGTTTCG	120			
Dd	61	CCGAGGGGTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGTGTGAAGGGGCAGCGTTTCG	120			
Qy	121	ACGTGGGCCCGCGCTACACGCACTTCAGATACCATCGCGGAGGGCGGTACGGGATGGTCA	180			
Dd	121	ACGTGGGCCCGCGCTACACGCACTTCAGATACCATCGCGGAGGGCGGTACGGGATGGTCA	180			
Qy	181	GCTCGGCCTATGACACGCTGCGCAAGACTCGCGTGCGCATCAAGAAGATCAGGCCCTTCG	240			
Dd	181	GCTCGGCCTATGACACGCTGCGCAAGACTCGCGTGCGCATCAAGAAGATCAGGCCCTTCG	240			
Qy	241	AACATCAGACCTACTGCCAGCGCACGCTCCGGGAGATCCAGATCTGTGCTGCTTCGGCC	300			
Dd	241	AACATCAGACCTACTGCCAGCGCACGCTCCGGGAGATCCAGATCTGTGCTGCTTCGGCC	300			
Qy	301	ATGAGATGTCAATCGGATCCGAGACAATTCTGCGGGGTCCAACCTGGAAGCCATGAGAG	360			
Dd	301	ATGAGAATGTCAATCGGATCCGAGACAATTCTGCGGGGTCCAACCTGGAAGCCATGAGAG	360			
Qy	361	ATGTCTACATTTGTGCAAGACTCGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGC	420			
Dd	361	ATGTCTACATTTGTGCAAGACTCGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGC	420			
Qy	421	AGTGAAGCAATACCATAATCTGCTACTTCTCTPACAGATCTCTGCGGGGCTCAAAGTACA	480			
Dd	421	AGTGAAGCAATACCATAATCTGCTACTTCTCTPACAGATCTCTGCGGGGCTCAAAGTACA	480			
Qy	481	TCACTTCGGCCACAGTGTCTCAACGAGATCTAAGCCCTCAACCTGCTCATCAACACCA	540			
Dd	481	TCCACTTCGGCCACAGTGTCTCAACGAGATCTAAGCCCTCAACCTGCTCATCAACACCA	540			
Qy	541	CCTGCGACCTTAAGATTGTGATTGTGGCTGGCCGGGATTCGGATCTCTGAGCATGACC	600			
Dd	541	CCTGCGACCTTAAGATTGTGATTGTGGCTGGCCGGGATTCGGATCTCTGAGCATGACC	600			
Qy	601	ACACCGGCTTCCTGACGGAGTATGTGGCTACGCGCTGGTACCGGGCCCCAGAGATCATGC	660			
Dd	601	ACACCGGCTTCCTGACGGAGTATGTGGCTACGCGCTGGTACCGGGCCCCAGAGATCATGC	660			
Qy	661	TGAATCCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTCTGGCTG	720			
Dd	661	TGAATCCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTCTGGCTG	720			
Qy	721	AGATGCTCTCTAACCGSCCATCTTCCTCGCAAGCACTACCTGGATCAGCTCAACCCACA	780			
Dd	721	AGATGCTCTCTAACCGSCCATCTTCCTCGCAAGCACTACCTGGATCAGCTCAACCCACA	780			
Qy	781	TTCTGGGCATCTCGGCTCCCCATCCCAGGAGCACTGAATGTGATCATCAACATGAAGG	840			
Dd	781	TTCTGGGCATCTCGGCTCCCCATCCCAGGAGCACTGAATGTGATCATCAACATGAAGG	840			
Qy	841	CCGAAACCTACCTACAGTCTCTGCCCCTCAAGACCAAGGTGGCTTGGGCCAAGCTTTTC	900			
Dd	841	CCGAAACCTACCTACAGTCTCTGCCCCTCAAGACCAAGGTGGCTTGGGCCAAGCTTTTC	900			
Qy	901	CCAAGTCAGACTCCAAAGCCCTTGACTGCTGGACCGGATGTTAACTTTAAACCCCAATA	960			
Dd	901	CCAAGTCAGACTCCAAAGCCCTTGACTGCTGGACCGGATGTTAACTTTAAACCCCAATA	960			
Qy	961	AACGGATCAAGTGGAGAAAGCGTGSCTCACCCCTACCTGGAGCAGTATATGACCCCGA	1020			
Dd	961	AACGGATCA-----	969			
Qy	1021	CGGATGACGATGGCGGAGGAGCCCTTCACTTCGCGATGGAGCTGGATGACCTACCTA	1080			

Db	970	-----CAGTGGCCGAGGAGCCCTTACCTTCGCCCATGTGGAGCTGGATGCACTACCTTA	1021
Qy	1081	AGGAGCGGCTGAAGAGCTCATCTTCAGGAGAGACAGCACGCTTCCAGCCCGGAGTGTCTGG	1140
Db	1021	AGGAGCGGCTGAAGGAGCTCATCTTCAGGAGAGACAGCACGCTTCCAGCCCGGAGTGTCTGG	1080
Qy	1141	AGGCCCCCTAGCCGAGACAGACATCTCTGTGCAACCTGGGGGCTGGAAACAGAACTGGCAAG	1200
Db	1081	AGGCCCCCTAGCCGAGACAGACATCTCTGTGCAACCTGGGGGCTGGAAACAGAACTGGCAAG	1140
Qy	1201	AGGCAAGAGGTCACTGAGGGGCTCTGTCAACGAGGACTGTGGCTCTGGCTTGCCTCTCTCC	1260
Db	1141	AGGCAAGAGGTCACTGAGGGGCTCTGTCAACGAGGACTGTGGCTCTGGCTTGCCTCTCTCC	1200
Qy	1261	CGCCAGACTGTTAGAAAATGGACACTGTGGCCAGCCCGGAGCTTGGCAGGCCGAGGCCGGG	1320
Db	1201	CGCCAGACTGTTAGAAAATGGACACTGTGGCCAGCCCGGAGCTTGGCAGGCCGAGGCCGGG	1260
Qy	1321	GTGGAGCATGGGCGCTGGGCCACTCTCTCTCTTTGTGTAGGGCTCTCAGGCTTCAGGCGAGCCCA	1380
Db	1261	GTGGAGCATGGGCGCTGGGCCACTCTCTCTCTTTGTGTAGGGCTCTCAGGCTTCAGGCGAGCCCA	1320
Qy	1381	AGGCCCTTCTCTCTCCCCACCCGCGCTCCGCCAGCGGGCTTCGGGAGCTCAGGTGGCCCCAGT	1440
Db	1321	AGGCCCTTCTCTCTCCCCACCCGCGCTCCGCCAGCGGGCTTCGGGAGCTCAGGTGGCCCCAGT	1380
Qy	1441	TCAATCTCCGCTGCTGTGTGGCGCTTACCTTCCGACGGTCCAGTCTCTTGGCAGT	1500
Db	1381	TCAATCTCCGCTGCTGTGTGGCGCTTACCTTCCGACGGTCCAGTCTCTTGGCAGT	1440
Qy	1501	TCTGGAAATGGAAGGTTCTGGCTGCCCCCAACCTGCTGAAGGGCAGAGGTGGAGGTTGGG	1560
Db	1441	TCTGGAAATGGAAGGTTCTGGCTGCCCCCAACCTGCTGAAGGGCAGAGGTGGAGGTTGGG	1500
Qy	1561	GGCCCTGAGTAGGAGACTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCACCTT	1620
Db	1501	GGCCCTGAGTAGGAGACTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCACCTT	1560
Qy	1621	AGTTTTCCTGAAGGAAATTCCTTAGTCTCAAAGGGCTAGCATCCCTCAGAGAGCAGGCCG	1680
Db	1561	AGTTTTCCTGAAGGAAATTCCTTAGTCTCAAAGGGCTAGCATCCCTCAGAGAGCAGGCCG	1620
Qy	1681	GGCCGAATCCCTCTCCCTGTCAAAGCTGTCACTTCGCGTGGCCCTCGCTGCTTGTGTGTG	1740
Db	1621	GGCCGAATCCCTCTCCCTGTCAAAGCTGTCACTTCGCGTGGCCCTCGCTGCTTGTGTGTG	1680
Qy	1741	GTGAGCAGAACTGGAGCTGGGGGCGGTGGAGAGCCCGCGGCCCTGTGCCACTCTCCTGACC	1800
Db	1681	GTGAGCAGAACTGGAGCTGGGGGCGGTGGAGAGCCCGCGGCCCTGTGCCACTCTCCTGACC	1740
Qy	1801	CGTCTAATATATAAATATAGAGATGTGTCTATGGCTG	1837
Db	1741	CGTCTAATATATAAATATAGAGATGTGTCTATGGCTG	1777

1805	DB	GGAGTGGAGATGGCGGCGCGCGCTCAGGGGGCGGGGCGGGGAGCCCCGTAGAACCC	1745
63	QY	GAGGGGTCCGCGCGGGGTCCCGGGGAGGTGAGATGGTGAAGGGGACGCGTTCCAGC	122
1745	DB	GAGGGGTTCGCCCCCGGGGTCCCCCGGAGAGTGGAGATCGTGAAGGGGACGCGTTCCAGC	1688
123	QY	GTGGGCGCGGCTACAACGAGTTGCAGTACATCGGGGAGGGCGGCTACGGCATGGTCCAGC	182
1685	DB	GTGGGCGCGGCTACAACGAGTTGCAGTACATCGGGGAGGGCGGCTACGGCATGGTCCAGC	162
183	QY	TCGGCTTATGACACACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCGAA	242
1625	DB	TCGGCTTATGACACACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCGAA	1566
243	QY	CATCAGACCTACTCCGACGGCACCGCTCCGGGAGATCCAGATCCCTGCTGCGCTCCCGCCAT	302
1565	DB	CATCAGACCTACTCCGACGGCACCGCTCCGGGAGATCCAGATCCCTGCTGCGCTCCCGCCAT	1506

QY 303 GAGATGTCATCGGCGCTCCGAGACATTTCTGCGGGCTCCACCTCGAAGCCATGAGAGAT 362
Db |||||
1505 GAGAAATGTCATCGGCATCCGAGACATTTCTGCGGGCTCCACCTCGAAGCCATGAGAGAT 1446
QY 363 GTCTACATTTGTGAGGACCTGATGAGAGACTGCTGTACAAAGTTGCTGAAAAGCCAGCAG 422
Db |||||
1445 GTCTACATTTGTGAGGACCTGATGAGAGACTGACCTGTACAAAGTTGCTGAAAAGCCAGCAG 1386
QY 423 CTGAGCAATGACATATCTGCTACTTCTCTTACAGATCTGCGGGGCTCAAGTACATC 482
Db |||||
1385 CTGAGCAATGACATATCTGCTACTTCTCTTACAGATCTGCGGGGCTCAAGTACATC 1326
QY 483 CACTCCGCGCAACGTGCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACACC 542
Db |||||
1325 CACTCCGCGCAACGTGCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACACC 1266
QY 543 TGGCACTTTAAGATTTGTGATTTTGGCTTGGCCCGGATTCGCCATCTCTGAGCATGACCAC 602
Db |||||
1265 TGGCACTTTAAGATTTGTGATTTTGGCTTGGCCCGGATTCGCCATCTCTGAGCATGACCAC 1206
QY 603 ACCGGCTTCTGACGAGTATGTGGCTACGCGTGTACGGGCCCCAGAGATCATGCTG 662
Db |||||
1205 ACCGGCTTCTGACGAGTATGTGGCTACGCGTGTACGGGCCCCAGAGATCATGCTG 1146
QY 663 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTTGGCTGAG 722
Db |||||
1145 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTTGGCTGAG 1086
QY 723 ATGCTCTCTAAACCGGGCCCATCTTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 782
Db |||||
1085 ATGCTCTCTAAACCGGGCCCATCTTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 1026
QY 783 CTGGGATCTCTGGGCTCCCATCCAGAGGACCTGAATTTGATCATCAACATGAAGGCC 842
Db |||||
1025 CTGGGATCTCTGGGCTCCCATCCAGAGGACCTGAATTTGATCATCAACATGAAGGCC 966
QY 843 CGAAACTTACCTACAGTCTCGCCCTCCAGACCAAGTGGCTTGGGCCCAAGCTTTTCCCC 902
Db |||||
965 CGAAACTTACCTACAGTCTCGCCCTCCAGACCAAGTGGCTTGGGCCCAAGCTTTTCCCC 906
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGTGGACCGGATGTTAACCTTTAACCCCAATAAA 962
Db |||||
905 AAGTCAGACTCCAAAGCCCTTGACCTGTGGACCGGATGTTAACCTTTAACCCCAATAAA 846
QY 963 CGGATCACAGTGAAGAGGCTTGGCTACCCCTTACTCTGAGCAGTACTATGACCCGAGC 1022
Db |||||
845 CGGATCACAGTGAAGAGGCTTGGCTACCCCTTACTCTGAGCAGTACTATGACCCGAGC 786
QY 1023 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTTACCTAAG 1082
Db |||||
785 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTTACCTAAG 726
QY 1083 GAGCGGCTGAAGGAGCTCATCTTCTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGGAG 1142
Db |||||
725 GAGCGGCTGAAGGAGCTCATCTTCTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGGAG 666
QY 1143 GCGCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCTTGGAAAGAACTGGCAAGAG 1202
Db |||||
665 GCGCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCTTGGAAAGAACTGGCAAGAG 626
QY 1203 GCAAGAGGTCACTGAGGGCTCTGTGTACCCAGGACCTGCTCTGCTGCTGCCCTCTCCCG 1262
Db |||||
625 -----GACCTTGCCTCTGCTGCTGCCCTCTCCCG 597
QY 1263 CCAGACTGTTAGAAAATGGAACACTGTGCCAGCCCGGACCTTGGCAGCCCGCGGGT 1322
Db |||||
596 CCAGACTGTTAGAAAATGGAACACTGTGCCAGCCCGGACCTTGGCAGCCCGCGGGT 537
QY 1323 GGAGCATGGGCTGGGCACTCTCTCTTGTGAGGCTTCCAGCTTCAGGC-----AG 1376
Db |||||
536 GGAGCATGGGCTGGGCACTCTCTCTTGTGAGGCTTCCAGCTTCAGGCAGGCGCAAG 477
QY 1377 GCCAAGGCTTCTCTCTCCCGCCAGCCGCTCTCCCGAGGCTCGGGGCTCGAGGTGGCC 1436

Db |||||
476 GCCAAGGCTTCTCTCTCCCGCCACCGCTCCCGACGGGGCTCGGAGCTCAGGTGGCC 417
QY 1437 CAGTTCAATCTCCC---GCTGCTGCTGCTGGCCCTTACCTTCCCAGGCTCCAGTCTC 1493
Db |||||
416 CAGTTCAATCTCCCCTGCTGCTGCTGGCCCTTACCTTCCCAGGCTCCAGTCTC 357
QY 1494 TGGCAGTTCTGGAATGGAAGGGTTCTGCTGCCCCCAACTGCTGAAGGCGAGGTGGAG 1553
Db |||||
356 TGGCAGTTCTGGAATGGAAGGGTTCTGCTGCCCCCAACTGCTGAAGGCGAGGTGGAG 297
QY 1554 GGTGGGGGGCGCTGAGTAGGGGACTCAGGGCCCATGCTGCCCCCTCATCTCATTCATAACC 1613
Db |||||
296 GGTGGGGGGCGCTGAGTAGGGGACTCAGGGCCCATGCTGCCCCCTCATCTCATTCATAACC 237
QY 1614 CCACCTAGTTTCCCTGAGGAGACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGC 1673
Db |||||
236 CCACCTAGTTTCCCTGAGGAGACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGC 177
QY 1674 CAGGCGGGCGGAATCCCTCTCTCAAAAGCTGTCACTTCGCGTGGCTCGCTGCTTCT 1733
Db |||||
176 CAGGCGGGCGGAAATCCCTCTCTCTCAAAAGCTGTCACTTCGCGTGGCTCGCTGCTTCT 117
QY 1734 GTGCTGTGTGAGCAGAAGTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTGCCACCTC 1793
Db |||||
116 GTGCTGTGTGAGCAGAAGTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTGCCACCTC 57
QY 1794 CTTGACCCGCTTAATATATATATATATAGAGATGTGTCTATGGCTG 1837
Db |||||
56 CTTGACCCGCTTAATATATATATATATAGAGATGTGTCTATGGCTG 13

RESULT 11
AA157910
ID AA157910 standard; cDNA; 1869 BP.
XX
AC AA157910;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 113.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00682191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
(HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
PI Zhou P, Goodrich R, Drmanac RT;
XX

Qy	1733	TGTGTGTGTGAGCAGAACTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTGCCACCT	1792		
Db	1753	TGTGTGTGTGAGCAGAACTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTGCCACCT	1812		
Qy	1793	CCCTGACCGCTCTAATATATAATATATAGAGATGTGTCTATGGCTG	1837		
Db	1813	CCCTGACCGCTCTAATATATAATATATAGAGATGTGTCTATGGCTG	1857		
RESULT 12					
ID	ADW12902				
XX	ADW12902 standard; cDNA; 1726 BP.				
AC	ADW12902;				
DT	07-APR-2005 (first entry)				
XX	Human SNAPK3V2 variant gene.				
DE					
XX	DNA purification; diagnosis; cancer; mitogen-activated protein kinase;				
KW	gene; ss.				
XX	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	12..1091			
FT	/*tag= a				
FT	/product= "SNAPK3V2 protein"				
PN	US2005013817-A1.				
XX					
PD	20-JAN-2005.				
XX					
PF	18-JUL-2003; 2003US-00623108.				
XX					
PR	18-JUL-2003; 2003US-00623108.				
XX					
PA	(DAIK/) DAI K.				
XX					
PI	Dai K;				
DR	WPI; 2005-080923/09.				
DR	P-PSDB; ADW12903.				
XX		New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing diseases, e.g. Cancers, associated with the deficiency of the SNAPK3 gene in a mammal.			
PT					
PT					
PT					
XX	Claim 4; SEQ ID NO 3; 55pp; English.				
XX		The invention relates to novel isolated polypeptide (I) comprising an amino acid sequence selected from sequences comprising 335 or 359 amino acids (ADW12901 or ADW12903) or its fragments. The polypeptides and CC polynucleotides are useful for diagnosing diseases, e.g. cancers, CC associated with the deficiency of the SNAPK3 gene in a mammal. The CC fragments of the polypeptides and polynucleotides can also be used as CC primers or probes. This sequence corresponds to the cDNA encoding the CC SNAPK3V1 variant protein.			
XX					
SQ	Sequence 1726 BP; 339 A; 562 C; 491 G; 334 T; 0 U; 0 Other;				
Query Match				86.8%; Score 1595; DB 14; Length 1726;	
Best Local Similarity				94.0%; Pred. No. 0;	
Matches 1726; Conservative				0; Mismatches 0; Indels 111; Gaps 2;	
Qy	1			GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGGAGCCCCGTAGAA	60
Db	1			GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGGAGCCCCGTAGAA	60
Qy	61			CCGAGGGGGTCCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGCCGTTCCG	120
Db	61	CCGAGGGGGTCCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGCCGTTCCG	120		

QY	121	ACGTGGCGCCGCTACACGAGTTGAGTACATCGCGAGGGCGCGTACGCGATGGTCA	180
DB	121	ACGTGGCGCCGCTACACGAGTTGAGTACATCGCGAGGGCGCGTACGCGATGGTCA	180
QY	181	GCTGGGCTATGACCAACGTTGGGCAAGACTCGGTTGGCCATCAAGAAGATCAGCCCTTCG	240
DB	181	GCTGGGCTATGACCAACGTTGGGCAAGACTCGGTTGGCCATCAAGAAGATCAGCCCTTCG	240
QY	241	RACATCAGACCTACTGCCAGCGCACGCTCCGGAGATCCAGATCTCTGCTGGCTTCCGCC	300
DB	241	RACATCAGACCTACTGCCAGCGCACGCTCCGGAGATCCAGATCTCTGCTGGCTTCCGCC	300
QY	301	ATGAGAAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTTGAAGCATAGAG	360
DB	301	ATGAGAAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTTGAAGCATAGAG	360
QY	361	ATGCTCTAATTGTGAGGACCTGTGAGAGATGAGACTGTGTAAGTTGCTGAAAGGCCAGC	420
DB	361	ATGCTCTAATTGTGAGGACCTGTGAGAGATGAGACTGTGTAAGTTGCTGAAAGGCCAGC	420
QY	421	AGCTGAGCAATGACCATATCTGCTACTTCTTACAGATCTCGGGGCGCTCAAGTACA	480
DB	421	AGCTGAGCAATGACCATATCTGCTACTTCTTACAGATCTCGGGGCGCTCAAGTACA	480
QY	481	TCCACTCCGCGCAACGTTGATTTGCGGCTGGCCCGGATTCGCGATCTGAGCATGACC	540
DB	481	TCCACTCCGCGCAACGTTGATTTGCGGCTGGCCCGGATTCGCGATCTGAGCATGACC	540
QY	541	CCTGCGACCTTAAAGATTTGATTTGCGGCTGGCCCGGATTCGCGATCTGAGCATGACC	600
DB	541	CCTGCGACCTTAAAGATTTGATTTGCGGCTGGCCCGGATTCGCGATCTGAGCATGACC	600
QY	601	ACACCGGCTTCTGAGCGAGTATGTTGCTACCGCTGGTACCGGGCCCGAGAGATCATGC	660
DB	601	ACACCGGCTTCTGAGCGAGTATGTTGCTACCGCTGGTACCGGGCCCGAGAGATCATGC	660
QY	661	TGAACCTCCAAGGGCTATACCAAGTCCATCGACATCTGTTGTTGGGCTGCAATCTGGCTG	720
DB	661	TGAACCTCCAAGGGCTATACCAAGTCCATCGACATCTGTTGTTGGGCTGCAATCTGGCTG	720
QY	721	AGATGCTCTTAACCGGGCCATCTTCCCTGGGCAAGCACTACCTGGATCAGCTCAACCA	780
DB	721	AGATGCTCTTAACCGGGCCATCTTCCCTGGGCAAGCACTACCTGGATCAGCTCAACCA	780
QY	781	TTCTGGGATCTCGGGCTCCCATCCAGAGGACCTGAATTTGATATCATCAACCAAGG	840
DB	781	TTCTGGGATCTCGGGCTCCCATCCAGAGGACCTGAATTTGATATCATCAACCAAGG	840
QY	841	CCCGAAACTACCTACAGTCTCTGCCCTCCAAGCAAGGTGGCTTGGGCAAGCTTTTCC	900
DB	841	CCCGAAACTACCTACAGTCTCTGCCCTCCAAGCAAGGTGGCTTGGGCAAGCTTTTCC	900
QY	901	CGAAGTCAGACTCCAAAGCCCTTGACTGTGACCGGATGTTAACTTTAACCCCAATA	960
DB	901	CGAAGTCAGACTCCAAAGCCCTTGACTGTGACCGGATGTTAACTTTAACCCCAATA	960
QY	961	AAAGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGCCGA	1020
DB	961	AAAGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGCCGA	969
QY	1021	CGGATGAGCCAGTGGCCGAGGAGCCCTTCCACTTCCCATGGAGCTGGATGACCTACCTA	1080
DB	970	-----CAGTGGCCGAGGAGCCCTTCCACTTCCCATGGAGCTGGATGACCTACCTA	1020
QY	1081	AGGAGCGGCTGAAGAGGCTCATCTTCCAGGACACGACGCTTCCAGCCCGGAGTGTCTGG	1140
DB	1021	AGGAGCGGCTGAAGAGGCTCATCTTCCAGGACACGACGCTTCCAGCCCGGAGTGTCTGG	1080
QY	1141	AGGCCCCCTTAGCCAGACAGATCTCTGCACTTGGGGCCCTGGGAACAGAACTGGCAAG	1200
DB	1081	AGGCCCCCTTAGCCAGACAGATCTCTGCACTTGGGGCCCTGGGAACAGAACTGGCAAG	1122
QY	1201	AGGCAAGAGGTCACTGAGGGGCTCTGTCTACCCAGGACCTTGTCTCTGCTGCCCCCTCTCC	1260

Db 661 TGAATCCAAAGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTCCTGGCTG 720
Qy 721 AGATGCTCTTAACCCGGCCCATCTTCCCTGGGAGCACTACCTGGATCAGCTCAACCA 780
Db 721 AGATGCTCTTAACCCGGCCCATCTTCCCTGGGAGCACTACCTGGATCAGCTCAACCA 780
Qy 781 TTCTGGGCATCCTGGGCTCCCATCCAGGAGGACCTGAATTGTATCATCAACATGAAG 840
Db 781 TTCTG----- 785
Qy 841 CCCGAACTACTACAGTCTCTGCCCCCTCCAGAACCAAGGTGGCTTGGGCCAAGCTTTTCC 900
Db 786 ----- 785
Qy 901 CCAAGTCAGATCCAAAGCCCTTGACCTGCTGACCGGATGTTAACTTTTACCCCAATA 960
Db 786 -----GCCCTTGACCTGCTGGACCGGATGTTAACTTTTACCCCAATA 828
Qy 961 AACGGATCAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Db 829 AACGGATCAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 888
Qy 1021 CGGATGAGCCAGTGGCCGAGGAGCCCTTTCACCTTCGCCATGGAGCTGGATGACCTACTA 1080
Db 889 CGGATGAGCCAGTGGCCGAGGAGCCCTTTCACCTTCGCCATGGAGCTGGATGACCTACTA 948
Qy 1081 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCGGAGTGTCTGG 1140
Db 949 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCGGAGTGTCTGG 1008
Qy 1141 AGGCCCCCTAGCCAGACAGATCTTGCACCCCTGGGGCTTGGAAACAGACTGGCAAG 1200
Db 1009 AGGCCCCCTAGCCAGACAGATCTTGCACCCCTGGGGCT----- 1050
Qy 1201 AGGCAAGAGGTCACTGAGGSCCTCTGTCAACCCAGGACCTGCTCTGCTGCCCCCTCTCC 1260
Db 1051 -----GGACCTGCTCTGCTGCTGCTGCTCTGCTGCTGCTCTGCTGCTGCTCTGCT 1077
Qy 1261 CGCCAGACTGTTAGAAATGGACATGTCGCCAGCCGGACCTTGGACGCCAGGCGGG 1320
Db 1078 CGCCAGACTGTTAGAAATGGACATGTCGCCAGCCGGACCTTGGACGCCAGGCGGG 1137
Qy 1321 GTGGAGCATGGGCTGCGCACCTCTCTCTCTTCTGAGGCTCCAGCTTCAGGAGGCCA 1380
Db 1138 GTGGAGCATGGGCTGCGCACCTCTCTCTCTTCTGAGGCTCCAGCTTCAGGAGGCCA 1197
Qy 1381 AGGCTTCTCTCCACACCGCCCTCCCAACCGGGCTCGGGAGCTCAGGTGGCCCACT 1440
Db 1198 AGGCTTCTCTCCACACCGCCCTCCCAACCGGGCTCGGGAGCTCAGGTGGCCCACT 1257
Qy 1441 TCAATCTCCGCTGCTGCTGCGGCTTACCTTCCAGCGCTCCAGTCTCTGCGACT 1500
Db 1258 TCAATCTCCGCTGCTGCTGCGGCTTACCTTCCAGCGCTCCAGTCTCTGCGACT 1317
Qy 1501 TCTGGAATGAAGGGTCTGCGTCCCACTGCTGAAGGGCAGAGGTGGAGGGTGGG 1560
Db 1318 TCTGGAATGAAGGGTCTGCGTCCCACTGCTGAAGGGCAGAGGTGGAGGGTGGG 1377
Qy 1561 GCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCATTAACACCCCACT 1620
Db 1378 GCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCATTAACACCCCACT 1437
Qy 1621 AGTTTCCCTGAAGAACATTCCTTAGTCTCAAGGGCTAGCTAGCTTCCCTGAGGAGCGGCG 1680
Db 1438 AGTTTCCCTGAAGAACATTCCTTAGTCTCAAGGGCTAGCTAGCTTCCCTGAGGAGCGGCG 1497
Qy 1681 GCGCGAATCCCTCTCCCTGTCAAAAGCTGTCTCACTTTCGGTGGCTTCTGCTGCTGCTG 1740
Db 1498 GCGCGAATCCCTCTCCCTGTCAAAAGCTGTCTCACTTTCGGTGGCTTCTGCTGCTGCTG 1557
Qy 1741 GTGAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGGCGCCCTCTGCACTTCCCTGACC 1800
Db 1558 GTGAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGGCGCCCTCTGCACTTCCCTGACC 1617

Qy 1801 CGTCTAATATATAAATATAGAGATGCTCTATGGCTG 1837
Db 1618 CGTCTAATATATAAATATAGAGATGCTCTATGGCTG 1654
RESULT 14
ADH48413
ID ADH48413 standard; DNA; 1699 BP.
XX ADH48413;
XX
DT 25-MAR-2004 (first entry)
XX Human KXP encoding DNA SEQ ID NO:71.
DE human; KXP; kinase; phosphatase; enzyme; cytostatic; anorectic;
XX immunosuppressive; KXP-Antagonist; KXP-Agonist; gene therapy;
KW autoimmune disorder; obesity; cancer; ds; gene.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 25..1032
FT /*tag= a
FT /product= "KPP"
XX
XX WO2004001008-A2.
XX
PD 31-DEC-2003.
XX
XX 19-JUN-2003; 2003WO-US019660.
XX
XX 21-JUN-2002; 2002US-0390652P.
XX
XX 15-JUL-2002; 2002US-0396196P.
XX
XX (INCY-) INCYTE CORP.
XX
XX Marquis JP, Baughn MR, Tran UK, Hafalia AJA, Kable AE,
PI Emerling BM, Elliott VS, Lindquist EA, Richardson TW, Khare R;
PI Swarnakar A, Lee SY, Ramkumar J, Chawla NK, Becha SD, Mason PM;
PI Hawkins PR, Bulloch SA, Jin P, Bhatia U, Burrill JD, Lee S;
PI Blake JJ, Ho A, Zheng W;
XX
XX WPI: 2004-082489/08.
XX P-PSDB; ADH48364.
XX
XX New human kinases and phosphatases (KPP) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional KPP e.g., cancer.
XX
XX Claim 5; SEQ ID NO 71; 336pp; English.
XX
XX The invention relates to novel isolated human kinases and phosphatases
CC (KPP) polypeptides. A protein of the invention has cytostatic, anorectic,
CC and immunosuppressive activity, and acts as a KXP-Antagonist, or KXP-
CC Agonist. A polynucleotide of the invention may have a use in gene
CC therapy. The polypeptide is useful for preparing a composition for
CC diagnosing or treating a disease or condition associated with decreased
CC expression or overexpression of functional KPP e.g., autoimmune
CC disorders, obesity or cancer. The sequences shown in ADH48392-ADH48440
CC represent KPP proteins of the invention.
XX
XX Sequence 1699 BP; 348 A; 538 C; 494 G; 319 T; 0 U; 0 Other;

Query Match 78.2%; Score 1437; DB 12; Length 1699;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 0; Indels 187; Gaps 3;
Qy 1 GAGGAGTGAGATGGCGGCGGCGGCTCAGGGGGGGGGGGAGCCCGTAGAA 60
Db 14 GAGGAGTGAGATGGCGGCGGCGGCTCAGGGGGGGGGGGAGCCCGTAGAA 73

QY 61 CCAGGGGGTCCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGCGTTGC 120
DB 74 CCAGGGGGTCCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGCGTTGC 133
QY 121 ACCTGGGCGCGCTACACGAGTTGAGTACATCGGCGAGGGGCGGTACGGCATGCTCA 180
DB 134 ACCTGGGCGCGCTACACGAGTTGAGTACATCGGCGAGGGGCGGTACGGCATGCTCA 193
QY 181 GCTCGGCTATGACCAAGTGGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 240
DB 194 GCTCGGCTATGACCAAGTGGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 253
QY 241 AACATCAGACTACTGCTCCAGGCAAGTCCGCGAGATCCAGATCCTGCTCGGCTTCGCGC 300
DB 254 AACATCAGACTACTGCTCCAGGCAAGTCCGCGAGATCCAGATCCTGCTCGGCTTCGCGC 313
QY 301 ATGAGATGTCATCGGATCCGAGACATTCGCGGGGCTCCACCTGGAGAGCCATGAGAG 360
DB 314 ATGAGATGTCATCGGATCCGAGACATTCGCGGGGCTCCACCTGGAGAGCCATGAGAG 373
QY 361 ATGTCTACATTTGTCAGGACTGATGAGACTGACCTGTACAAAGTTGCTGAAAAGCCAGC 420
DB 374 ATGTCTACATTTGTCAGGACTGATGAGACTGACCTGTACAAAGTTGCTGAAAAGCCAGC 433
QY 421 AGCTGACCAATGACCATATCTGCTACTTCTCTACAGATCTCTGCGGGGCTCAAGTACA 480
DB 434 AGCTGACCAATGACCATATCTGCTACTTCTCTACAGATCTCTGCGGGGCTCAAGTACA 493
QY 481 TCCACTCCGCAAGTGTCTCCACGAGATCTTAAGCCCTCCAACTGCTCATCAACACCA 540
DB 494 TCCACTCCGCAAGTGTCTCCACGAGATCTTAAGCCCTCCAACTGCTCATCAACACCA 553
QY 541 CCTGCGACCTTAAGATTTGATTTCCGCGCTGGCCCGGATTTGCCGATCCTGAGCATGACC 600
DB 554 CCTGCGACCTTAAGATTTGATTTCCGCGCTGGCCCGGATTTGCCGATCCTGAGCATGACC 613
QY 601 ACACCGGCTTCTGACGAGATATGCTGCTACGCGCTGGTACCGGGCCCAAGATATGCG 660
DB 614 ACACCGGCTTCTGACGAGATATGCTGCTACGCGCTGGTACCGGGCCCAAGATATGCG 673
QY 661 TGAATCTCAAGGCTATACCAAGTCCATCGACATCTGCTGCTGGGCTGATTTCTGGCTG 720
DB 674 TGAATCTCAAGGCTATACCAAGTCCATCGACATCTGCTGCTGGGCTGATTTCTGGCTG 733
QY 721 AGATGCTCTTAACCGGCCCATCTTCCCTGCGAGCACTACTCGATCAGCTCAACCCACA 780
DB 734 AGATGCTCTTAACCGGCCCATCTTCCCTGCGAGCACTACTCGATCAGCTCAACCCACA 793
QY 781 TTCTGGGCTATCTGGGCTCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAAG 840
DB 794 TTCTG----- 798
QY 841 CCGAAGACTACCTACAGTCTCTGCCCTCCAGACCAAGTGGCTTGGGCCAGCTTTTCC 900
DB 799 ----- 798
QY 901 CCAAGTCAGACTCCAAAGCCCTTACCTGCTGGACCGGATGTTAACTTTAAACCCCAATA 960
DB 799 -----GCCCTTGAACCTTGTCTGGACCGGATGTTAACTTTAAACCCCAATA 841
QY 961 AACGGATCACAGTGGAGGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA 1020
DB 842 AACGGATCACAGTGGAGGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA 901
QY 1021 CGGATGAGCCAGTGGCGAGAGCCCTTCACTTGGCCATGGAGCTGGATGACCTACTTA 1080
DB 902 CGGATGAGCCAGTGGCGAGAGCCCTTCACTTGGCCATGGAGCTGGATGACCTACTTA 961
QY 1081 AGGAGCGCTCAAGGAGCTCATCTTCCAGGAGACAGACCGCTTCCAGCCCGGAGTGTGG 1140
DB 962 AGGAGCGCTCAAGGAGCTCATCTTCCAGGAGACAGACCGCTTCCAGCCCGGAGTGTGG 1021
QY 1141 AGGCCCTTACGCCAGACAGACATCTCTGCAACCTTGGGGCTGGAAACAGAACTGGCAAG 1200

DB 1022 AGGCCCTTACGCCAGACAGACATCTCTGACCCCTGGGGCT----- 1063
QY 1201 AGGCAAGAGGTCACTGAGGGCTCTGTCAACCAAGGACCTGCTCTCTGCTGCCCTCTCC 1260
DB 1064 -----GGACCTGCTCTCTGCTGCCCTCTCC 1090
QY 1261 CGCAGACTGTTAGAAAATGGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGGCGGG 1320
DB 1091 CGCAGACTGTTAGAAAATGGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGGCGGG 1150
QY 1321 GTGGAGCATGGGCTCGGCCACCTCTCTCTTCTGAGGCTCCAGCTTCAGGCGAGGCCA 1380
DB 1151 GTGGAGCATGGGCTCGGCCACCTCTCTCTTCTGAGGCTCCAGCTTCAGGCGAGGCCA 1210
QY 1381 AGGCCCTTCTCTCCCAACCGCCCTCCCAACCGGGGCTCGGGAGCTCAGGTGCCCCAGT 1440
DB 1211 AGGCCCTTCTCTCCCAACCGCCCTCCCAACCGGGGCTCGGGAGCTCAGGTGCCCCAGT 1270
QY 1441 TCAATCTCCGCT-----GCTGCTGCTGGGCTCTTACTTCCCAGAGTCCAGTCTCTGG 1496
DB 1271 TCAATCTCCGCTGCTGGCTGCTGCTGCGCCCTTACTTCCCAGAGTCCAGTCTCTGG 1330
QY 1497 CAGTCTCGAATCGAAGGGTCTGCGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGT 1556
DB 1331 CAGTCTCGAATCGAAGGGTCTGCGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGT 1390
QY 1557 GGGGGCGCTGAGTAGGGACTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCCA 1616
DB 1391 GGGGGCGCTGAGTAGGGACTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCCA 1450
QY 1617 CCTAGTCTTCCCTGAAGGAACATCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCAG 1676
DB 1451 CCTAGTCTTCCCTGAAGGAACATCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCAG 1510
QY 1677 GCCGGCCGAATCCCTCCCTGTCAAAGCTGTCACCTTGGGCTGCCCTGCTCTCTGTG 1736
DB 1511 GCCGGCCGAATCCCTCCCTGTCAAAGCTGTCACCTTGGGCTGCCCTGCTCTCTGTG 1570
QY 1737 TGTGCTGAGCAGAAGTGGAGCTGGGGGCTGGAGAGCCCGGCGCCCTGCCACCTCCCT 1796
DB 1571 TGTGCTGAGCAGAAGTGGAGCTGGGGGCTGGAGAGCCCGGCGCCCTGCCACCTCCCT 1630
QY 1797 GACCGCTTAATATATAATATATAGAGATGTCTATGGCTG 1837
DB 1631 GACCGCTTAATATATAATATATAGAGATGTCTATGGCTG 1671

RESULT 15

ADH59631

ID ADH59631 standard; DNA; 1140 BP.

XX AC ADH59631;

XX DT 25-MAR-2004 (first entry)

XX DE Erk1 encoding sequence.

XX KW cardiac disease; Raf-1; MEK1; Cardiant; telethonin; &agr-actinin;
XX KW hypertrophy; ds; MEK2; Erk1; Erk2.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX CDS 1..1140

XX FT /*tag= a

XX FT /product= "Erk1"

XX WO2003025205-A2.

XX PN 27-MAR-2003.

XX PD 18-SEP-2002; 2002WO-BP010489.

XX PF

XX 19-SEP-2001; 2001US-0323566P.
PR 24-SEP-2001; 2001US-0324625P.
XX (MEDI-) MEDIGENE AG.
XX Nave B, Roenicke V, Leclair S, Funk M, Reuner B, Brinkmann K;
PI Henkel T;
XX WPI; 2003-371821/35.
DR P-PSDB; ADH59632.
XX
XX Identifying and/or obtaining a compound useful for preventing or treating
PT cardiac diseases, particularly congestive heart failure, comprises
PT quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
PT on the compound.
XX
XX Claim 33; SEQ ID NO 7; 41pp; English.
XX
CC The present invention relates to identifying and obtaining a compound
CC useful in the prevention or treatment of cardiac diseases, comprising
CC quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
CC on the compound. The method is useful for identifying and/or obtaining
CC compounds that may be used in the prevention or treatment of cardiac
CC diseases, particularly congestive heart failure. The compound or protein
CC is also used for the preparation of a pharmaceutical composition for
CC prevention or treatment of a disease related to hypertrophy or impaired
CC or increased activation of telethonin (T-Cap), α myosin, troponin, MHC, actin,
CC titin, myomesin, nebulin, tropomyosin, troponin, Erk1/2 and/or MLCK.
CC The protein or antibody which specifically recognizes the
CC activated/phosphorylated form of the above polypeptide, is used for the
CC preparation of a composition for diagnosing a disease or a predisposition
CC for a disease related to hypertrophy or related to impaired or increased
CC activation of Raf-1, MEK1/2 and/or Erk1/2. The present sequence
CC represents Erk1 encoding sequence.
XX
XX Sequence 1140 BP; 247 A; 357 C; 324 G; 212 T; 0 U; 0 Other;
XX
Query Match 62.0%; Score 1138.4; DB 10; Length 1140;
Best Local Similarity 99.9%; Pred. No. 2.2e-247;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX 12 ATGGCGCGCGCGCGCTCAGGGGGGCGGGGGGCGGAGCCCGTAGAACCCGAGGGGGTC 71
DB 1 ATGGCGCGCGCGCGCTCAGGGGGGCGGGGGGCGGAGCCCGTAGAACCCGAGGGGGTC 60
XX 72 GGCGCGGGGGTCCCGGGGGAGGTGGAGATGGTGAAGGGGCGCGGTCGACGTGGGCGCG 131
DB 61 GGCGCGGGGGTCCCGGGGGAGGTGGAGATGGTGAAGGGGCGCGGTCGACGTGGGCGCG 120
XX 132 CGGTACACGCGTTCGAGTACATCGGGGGGGCGGTAGCGGATGGTTCGGGCTAT 191
DB 121 CGGTACACGCGTTCGAGTACATCGGGGGGGCGGTAGCGGATGGTTCGGGCTAT 180
XX 192 GACACGTCGCGAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAACATCAGACC 251
DB 181 GACACGTCGCGAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAACATCAGACC 240
XX 252 TACTGCCAGCGCGCTCCGGGGAGATCCAGATCTCTGCTGGCTTCGGCCATGAGATGTC 311
DB 241 TACTGCCAGCGCGCTCCGGGGAGATCCAGATCTCTGCTGGCTTCGGCCATGAGATGTC 300
XX 312 ATCGGATCCGAGACATTCGCGGGGGTCCACCCCTGGAAGCCATGAGATGTCTACATT 371
DB 301 ATCGGATCCGAGACATTCGCGGGGGTCCACCCCTGGAAGCCATGAGATGTCTACATT 360
XX 372 GTGCGAGCTGATGGAGATGACCTGTACAGTTGCTGAAGAAGCAGCGCTGAGCAAT 431
DB 361 GTGCGAGCTGATGGAGATGACCTGTACAGTTGCTGAAGAAGCAGCGCTGAGCAAT 420
XX 432 GACCATATCTGCTACTTCCTTACAGATCTCTGGGGGGCTCAAGTACATCCACTCCGCC 491
DB 421 GACCATATCTGCTACTTCCTTACAGATCTCTGGGGGGCTCAAGTACATCCACTCCGCC 480

QY 492 AACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACTTGCACCTT 551
DB 481 AACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACTTGCACCTT 540
QY 552 AAGATTTGTGATTTCCGGCTTGGCCCGGATTCGCGATCTCTGAGCATGACACACCGGCTTC 611
DB 541 AAGATTTGTGATTTCCGGCTTGGCCCGGATTCGCGATCTCTGAGCATGACACACCGGCTTC 600
QY 612 CTGACGGAGTATGTGGCTACGGGCTGGTACCGGGGCCCGAGAGATCATGCTGAACCTCCAG 671
DB 601 CTGACGGAGTATGTGGCTACCGGCTGGTACCGGGGCCCGAGAGATCATGCTGAACCTCCAG 660
QY 672 GGCTATACCAAGTCCCATCGACATCTGTGCTGTGGGCTGCATCTCTGCTGAGATGCTCTCT 731
DB 661 GGCTATACCAAGTCCCATCGACATCTGTGCTGTGGGCTGCATCTCTGCTGAGATGCTCTCT 720
QY 732 AACCGGCGCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCACTTCTGGGCTTC 791
DB 721 AACCGGCGCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCACTTCTGGGCTTC 780
QY 792 CTGGGCTCCCATCCAGGAGGACCTGATTTGTATCATCAACATGAGGCGCGAACTAC 851
DB 781 CTGGGCTCCCATCCAGGAGGACCTGATTTGTATCATCAACATGAGGCGCGAACTAC 840
QY 852 CTACAGTCTCTGCCCTCCAAAGACCAAGGTGGCTTGGGCGCAAGCTTTTCCCAAGTCAGAC 911
DB 841 CTACAGTCTCTGCCCTCCAAAGACCAAGGTGGCTTGGGCGCAAGCTTTTCCCAAGTCAGAC 900
QY 912 TCCAAAGCCCTTGACCTCTGCGACCGGATGTTAACTTTAAACCCCAATAAACGGATCA 971
DB 901 TCCAAAGCCCTTGACCTCTGCGACCGGATGTTAACTTTAAACCCCAATAAACGGATCA 960
QY 972 GTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCGGAGATGAGCCA 1031
DB 961 GTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCGGAGATGAGCCA 1020
QY 1032 GTGGCGGAGGAGCCCTTCACTTCCCATGGAGCTGGATGACCTAACGAGGCGGCTG 1091
DB 1021 GTGGCGGAGGAGCCCTTCACTTCCCATGGAGCTGGATGACCTAACGAGGCGGCTG 1080
QY 1092 AAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTCTGAGAGCCGCCCTAG 1151
DB 1081 AAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTCTGAGAGCCGCCCTAG 1140

Search completed: February 7, 2006, 12:49:02
Job time : 1133.96 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:29:45 ; Search time 7228.32 Seconds
(without alignments)
11890.431 Million cell updates/sec

Title: US-10-623-108-5
Perfect score: 1837
Sequence: 1 gaggagtggagatggcggcg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_est7.*

8: gb_est8.*

9: gb_est9.*

10: gb_est10.*

11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1712.4	93.2	1775	4	CR620058 full-leng
2	1699.4	92.5	1762	4	CR621455 full-leng
3	1694.4	92.2	1764	4	CR606522 full-leng
4	1688.4	91.9	1751	4	CR595737 full-leng
5	1681.4	91.5	1751	4	CR597147 full-leng
6	1673.4	91.1	1854	4	CR603463 full-leng
7	1666.4	90.7	1729	4	CR621977 full-leng
8	1644.4	89.5	1707	4	CR621688 full-leng
9	1615.4	87.9	1678	4	CR600245 full-leng
10	1577.4	85.9	1718	4	CR596064 full-leng
11	1549.4	84.3	1612	4	CR607536 full-leng
12	1034.2	56.3	1130	1	AL530655 AL530655
13	1008	54.9	1201	1	AL530654 AL530980
14	960.2	52.3	1016	1	AL530980 AL530980
15	945.2	51.5	1108	5	EX445445 EX445445
16	934.8	50.9	1013	1	AL515713 AL515713
17	933.8	50.8	946	1	AL526297 AL526297
18	913.8	49.7	989	1	AL583198 AL583198
19	908.6	49.5	985	5	EX398410 EX398410
20	906	49.3	949	1	AL549418 AL549418
21	898.4	48.9	986	1	AL529832 AL529832
22	883.8	48.1	1089	3	BM548079 AGENCOURT

23	882.2	48.0	910	5	BX386989
24	876.8	47.7	898	1	AL559377
25	875	47.6	875	10	AY403905
26	872.6	47.5	1131	3	BM462994
27	867.2	47.2	902	1	AL533916
28	867	47.2	1083	5	BX333599
29	862.6	47.0	921	7	CR999816
30	859	46.8	1036	3	BQ66884
31	856.2	46.6	909	7	CR981539
32	855.2	46.6	882	5	BQ980149
33	854.6	46.5	1015	2	BI225074
34	853.2	46.4	963	1	AL526262
35	850.4	46.3	968	5	BX354177
36	845	46.0	877	1	AL558763
37	843	45.9	843	1	AL516647
38	842.8	45.9	847	10	AY403906
39	837.4	45.6	1053	3	BM468615
40	832.8	45.3	864	1	AL576621
41	830.4	45.2	1067	1	AL544810
42	829.6	45.2	971	3	BI768741
43	825.6	44.9	1128	1	AL583197
44	820.8	44.7	1080	5	EX445446
45	819.4	44.6	847	3	BI768732

ALIGNMENTS

RESULT 1
CR620058
LOCUS full-length cDNA clone CS0DA001YE10 of Neuroblastoma of Homo sapiens (human).
DEFINITION CR620058
ACCESSION CR620058.1 GI:50500865
VERSION HTC; CNSLT_cDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1775)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1775)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source
1..1775
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA001YE10"
/tissue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 93.2%; Score 1712.4; DB 4; Length 1775;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1774; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers source 1..1762 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D008YK04" /issue.type="Neuroblastoma Cot 50-normalized" /plasmid="pCMVSPORT_6"
ORIGIN	Query Match 92.5%; Score 1699.4; DB 4; Length 1762; Best Local Similarity 97.1%; Pred. No. 0; Matches 1761; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
Qy	11 GATGGCGGCGCGCGCTCAGGGGGGCGGGGGCGGGAGCCCCGTAGAACCAGAGGGGT 70
Db	1 GATGGCGGCGCGCGCGCTCAGGGGGGCGGGGGCGGGAGCCCCGTAGAACCAGAGGGGT 60
Qy	71 CGGCCCGGGGGTCCCGGGGGAGTTGAGATGTTGAAGGGGCAGCGTTTCGACGTGGGGCC 130
Db	61 CGGCCCGGGGGTCCCGGGGGAGTTGAGATGTTGAAGGGGCAGCGTTTCGACGTGGGGCC 120
Qy	131 GCCTTACACGCAGTTTCAGTACATCGCGGAGGCGCGGTACGGCATGCTCAGCTCGGCCA 190
Db	121 GCCTTACACGCAGTTTCAGTACATCGCGGAGGCGCGGTACGGCATGCTCAGCTCGGCCA 180
Qy	191 TGACCAAGTCGCGAAGATCTCGTGCGCATCAAGAAAGATCAGCCCCTTCGAACATCAGAC 250
Db	181 TGACCAAGTCGCGAAGATCTCGTGCGCATCAAGAAAGATCAGCCCCTTCGAACATCAGAC 240
Qy	251 CTACTCGCACGCGACCTCGGGAGATCCAGATCTCTGCTGGCGCTTCGCGCATGAGATGT 310
Db	241 CTACTCGCACGCGACCTCGGGAGATCCAGATCTCTGCTGGCGCTTCGCGCATGAGATGT 300
Qy	311 CATCGGCATCCGAGACAATTCTGGGCGCTCCACCTTGGAAAGCATGAGAGATGTCTACAT 370
Db	301 CATCGGCATCCGAGACAATTCTGGGCGCTCCACCTTGGAAAGCATGAGAGATGTCTACAT 360
Qy	371 TGTGCAAGAACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAAGCCAGCGCTGAGCAA 430
Db	361 TGTGCAAGAACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAAGCCAGCGCTGAGCAA 420
Qy	431 TGACCATATCTGTAATTCTTACAGATCTCGGGGGCTCAAGTATCATCTCCATCCGCGC 490
Db	421 TGACCATATCTGTAATTCTTACAGATCTCGGGGGCTCAAGTATCATCTCCATCCGCGC 480
Qy	491 CAACGTGCTCCACCGAGATCTAAAGCCCTCCAACCTGCTCATCAACACCACTCGGACCT 550
Db	481 CAACGTGCTCCACCGAGATCTAAAGCCCTCCAACCTGCTCATCAACACCACTCGGACCT 540
Qy	551 TAAGATTGTGATTTTCCGGCTGGCCCGGATTCGGATCTGAGCATGACCAACCGGGTTT 610
Db	541 TAAGATTGTGATTTTCCGGCTGGCCCGGATTCGGATCTGAGCATGACCAACCGGGTTT 600
Qy	611 CTTGACGGAGTATGTGGCTACGGCTGGTACCGGGCCCGAGAGATCATCTGAACTCCAA 670
Db	601 CTTGACGGAGTATGTGGCTACGGCTGGTACCGGGCCCGAGAGATCATCTGAACTCCAA 660
Qy	671 GGGCTATPACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTTCTGGCTGAGATCTCT 730
Db	661 GGGCTATPACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTTCTGGCTGAGATCTCT 720
Qy	731 TAACCGGCCCCATCTTCCCTGGCAAGCATCACTCGGATCAGCTCAACCAATTTCTGGGCAT 790

Db 1081 AGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGAGGCGCCCTAGCCCA 1140

Qy 1156 GACAGACATCTGCAACCTGGGGCTGGACAGACTGGCAAGAGGCAAGAGGTCACT 1215

Db 1141 GACAGACATCTGCAACCTGGGGCT----- 1167

Qy 1216 GAGGGCTCTGTACCCAGGACCTGCTCTGCTGCTCCCTCTCCCGCCAGACTGTTAGA 1275

Db 1168 -----GGACCTGCTCTGCTGCTCCCTCTCCCGCCAGACTGTTAGA 1209

Qy 1276 AAATGACACTGTGCCAGCCCGACCTTGGCAGCCAGCCGCGGTGAGCAATGGGCT 1335

Db 1210 AAATGACACTGTGCCAGCCCGACCTTGGCAGCCAGCCGCGGTGAGCAATGGGCT 1269

Qy 1336 GGCACTCTCTCTTGTGAGGCTTCCAGCTTCAAGGAGGCGCAGGCTTCTCTCC 1395

Db 1270 GGCACTCTCTCTTGTGAGGCTTCCAGCTTCAAGGAGGCGCAGGCTTCTCTCC 1329

Qy 1396 CACCCGCTCTCCCAAGGCGCTCGGAGCTCAGGTGGCCCAAGTTCAATCTCCCGCTGC 1455

Db 1330 CACCCGCTCTCCCAAGGCGCTCGGAGCTCAGGTGGCCCAAGTTCAATCTCCCGCTGC 1389

Qy 1456 TGCTGTGCGCCCTTACCTTCCCGAGCGTCCAGTCTCTGGCAGTTCTGGAATGGAAGG 1515

Db 1390 TGCTGTGCGCCCTTACCTTCCCGAGCGTCCAGTCTCTGGCAGTTCTGGAATGGAAGG 1449

Qy 1516 TTCTGCTGCGCCCAAGCTGTGAAGGCGAGAGGTGAGGCTGGGGCGCTGAGTAGGA 1575

Db 1450 TTCTGCTGCGCCCAAGCTGTGAAGGCGAGAGGTGAGGCTGGGGCGCTGAGTAGGA 1509

Qy 1576 CTGAGGCGCAATGCTGCGCCCTCATCTCATTTCAAAACCCCAAGTTTCCCTGAAGGA 1635

Db 1510 CTGAGGCGCAATGCTGCGCCCTCATCTCATTTCAAAACCCCAAGTTTCCCTGAAGGA 1569

Qy 1636 ACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCGCGGCGAATCCCTCC 1695

Db 1570 ACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCGCGGCGAATCCCTCC 1629

Qy 1696 CTGTCAGCTGTCACTTCGCGTGGCTCGCTCTCTGTTGTTGCTGAGCAGAGTGA 1755

Db 1630 CTGTCAGCTGTCACTTCGCGTGGCTCGCTCTCTGTTGTTGTTGAGCAGAGTGA 1689

Qy 1756 GCTGGGGGGCGTGAGAGCGCGCGCCCTGCAACCTCCCTGACCCGCTTAATATATAA 1815

Db 1690 GCTGGGGGGCGTGAGAGCGCGCGCCCTGCAACCTCCCTGACCCGCTTAATATATAA 1749

Qy 1816 TA 1817

Db 1750 TA 1751

RESULT 5

CR597147

LOCUS

DEFINITION

full-length cDNA clone CS0D013YD03 of T cells (Jurkat cell line)

CR597147 1751 bp mRNA linear HTC 21-JUL-2004

full-length cDNA clone CS0D013YD03 of T cells (Jurkat cell line)

CR597147

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 1751)

AUTHORS

Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/

Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1751)

Genoscope.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

Location/Qualifiers

source

1..1751

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D013YD03"

/tissue_type="T cells (Jurkat cell line) Cot

10-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 91.5%; Score 1681.4; DB 4; Length 1751;

Best Local Similarity 97.1%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

Qy 1 GAGGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAA 60

Db 8 GAGGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAA 67

Qy 61 CGAGGGGTTCGGCCCGGGGGTCCCGGGAGAGTGGAGATGGTGAAGGGGACCGTTGC 120

Db 68 CGAGGGGTTCGGCCCGGGGGTCCCGGGAGAGTGGAGATGGTGAAGGGGACCGTTGC 127

Qy 121 AGTGGGGCGCGCTACACGAGTTGCACTACATCGGGAGGCGCGTACGGCATGGTCA 180

Db 128 AGTGGGGCGCGCTACACGAGTTGCACTACATCGGGAGGCGCGTACGGCATGGTCA 187

Qy 181 GCTCGGCTTATGACCAACGTCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCG 240

Db 188 GCTCGGCTTATGACCAACGTCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCG 247

Qy 241 AACATCAGACCTTACCTCCAGCGCAGCTCCGGAGATCCAGATCTCTGCGCTTCCGCC 300

Db 248 AACATCAGACCTTACCTCCAGCGCAGCTCCGGAGATCCAGATCTCTGCGCTTCCGCC 307

Qy 301 ATGAGAAATGTCAATCGCATCCGAGATCTCTCGGGCGTCCACCTTGAAGCCATGAGAG 360

Db 308 ATGAGAAATGTCAATCGCATCCGAGATCTCTCGGGCGTCCACCTTGAAGCCATGAGAG 367

Qy 361 ATGTCTTACATTTGTCAGGACCTGATGGAGCTGACCTGTACAAAGTTGCTGAAAAGCCAGC 420

Db 368 ATGTCTTACATTTGTCAGGACCTGATGGAGCTGACCTGTACAAAGTTGCTGAAAAGCCAGC 427

Qy 421 AGCTGAGCAATGACCAATATCTGCTACTTCTTACCAAGATCTCTGCGGGGCTCAAGTACA 480

Db 428 AGCTGAGCAATGACCAATATCTGCTACTTCTTACCAAGATCTCTGCGGGGCTCAAGTACA 487

Qy 481 TCCACTTCGCGCAACGTCCTCCACGAGATCTAAAGCCCTCCCAACCTGTCTATCAACACCA 540

Db 488 TCCACTTCGCGCAACGTCCTCCACGAGATCTAAAGCCCTCCCAACCTGTCTATCAACACCA 547

Qy 541 CTTGCGACCTTAAAGATTTGTGATTTTCGGCTTCGGCCCGGATTCGGATCTGAGCATGACC 600

Db 548 CTTGCGACCTTAAAGATTTGTGATTTTCGGCTTCGGCCCGGATTCGGATCTGAGCATGACC 607

Qy 601 ACACGGCTTCTGACGGAGTATGCGCTACCGGTGTTACCGGGCCCCCAGAGATCATGC 660

Db 608 ACACGGCTTCTGACGGAGTATGCGCTACCGGTGTTACCGGTGTTACCGGGCCCCCAGAGATCATGC 667

Qy 661 TGAACCTCCAAAGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720

Db 668 TGAACCTCCAAAGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 727

Qy 721 AGATGCTCTTAACCGGGGCCATCTTTCCTGGCAAGCACTACCTGGATCAGCTCAACACCA 780

```
Db 728 AGATGCTCTTAACCGGCCCATCTTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 787
Qy 781 TTCTGGGCATCTGGGCTCCCAATCCAGAGAGACCTGTAATTGTATCATCAACATGAAG 840
Db 788 TTCTGGGCATCTGGGCTCCCAATCCAGAGAGACCTGTAATTGTATCATCAACATGAAG 847
Qy 841 CCCGAACTACTACAGTCTCTGCTCTCAAGACCAAGTGGCTTGGCCCAAGCTTTTC 900
Db 848 CCCGAACTACTACAGTCTCTGCTCTCAAGACCAAGTGGCTTGGCCCAAGCTTTTC 907
Qy 901 CCAAGTCAGACTCAAAAGCCCTTGAACCTGCTGAGACCGGATGTTAACTTTAACCCCAATA 960
Db 908 CCAAGTCAGACTCAAAAGCCCTTGAACCTGCTGAGACCGGATGTTAACTTTAACCCCAATA 967
Qy 961 AACGGATCAGAGTGGAGGAGCGCTGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Db 968 AACGGATCAGAGTGGAGGAGCGCTGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1027
Qy 1021 CGGATGAGCCAGTGGCCGAGGAGCCCTTTCACCTTTCGCCATGGAGCTGGATGACCTACTA 1080
Db 1028 CGGATGAGCCAGTGGCCGAGGAGCCCTTTCACCTTTCGCCATGGAGCTGGATGACCTACTA 1087
Qy 1081 AGGAGCGGCTGAAGAGCTCATCTTCCAGAGACAGACACGCTTCCAGCCCGGAGTGTGG 1140
Db 1088 AGGAGCGGCTGAAGAGCTCATCTTCCAGAGACAGACACGCTTCCAGCCCGGAGTGTGG 1147
Qy 1141 AGGCCCTTACCCAGACAGACATCTCTGACCTCTGGGCTTGGACAGAACTGGCAAG 1200
Db 1148 AGGCCCTTACCCAGACAGACATCTCTGACCTCTGGGCTTGGACAGAACTGGCAAG 1189
Qy 1201 AGGCAAGAGTCACTGAGGGGCTCTGTGACCCAGAGCTTGCCTCTGCTGCCCTCTCC 1260
Db 1190 -----GGACCTGCCCTCTGCTGCCCTCTCC 1216
Qy 1261 CGCCAGACTGTTAGAAATGGAACAATGTCGTCAGCCCGGAGCTTGGAGCCGAGCCGG 1320
Db 1217 CGCCAGACTGTTAGAAATGGAACAATGTCGTCAGCCCGGAGCTTGGAGCCGAGCCGG 1276
Qy 1321 GTGGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAAGGAGGCCA 1380
Db 1277 GTGGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAAGGAGGCCA 1336
Qy 1381 AGGCTTTCTCTCCCAACCCGCTTCCCAAGGAGCTTGGAGCTCAGGTGGCCCAAGT 1440
Db 1337 AGGCTTTCTCTCCCAACCCGCTTCCCAAGGAGCTTGGAGCTCAGGTGGCCCAAGT 1396
Qy 1441 TCAATCTCCGCTGCTGCTGCGCCCTTACCTTCCCAAGGCTTCCAGCTTCTTGGCAGT 1500
Db 1397 TCAATCTCCGCTGCTGCTGCGCCCTTACCTTCCCAAGGCTTCCAGCTTCTTGGCAGT 1456
Qy 1501 TCTGGAATGGAGGGTTCTGGCTGCGCCCAACCTGCTGGAAGGAGAGGTGGAGGGTGGG 1560
Db 1457 TCTGGAATGGAGGGTTCTGGCTGCGCCCAACCTGCTGGAAGGAGAGGTGGAGGGTGGG 1516
Qy 1561 GGCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAAACCCCACT 1620
Db 1517 GGCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAAACCCCACT 1576
Qy 1621 AGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATTCCTGAGGAGCCAGGCG 1680
Db 1577 AGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATTCCTGAGGAGCCAGGCG 1636
Qy 1681 GGCGAATCCCTTCCCTGTCAGAGTGTCACTTCCGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1637 GGCGAATCCCTTCCCTGTCAGAGTGTCACTTCCGCTGCTGCTGCTGCTGCTGCTG 1696
Qy 1741 GTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTGCGACCTCC 1795
Db 1697 GTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTGCGACCTCC 1751
```

RESULT 6

```
CR603463 1854 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DN005YA14 of Adult brain of Homo sapiens
DEFINITION (human).
ACCESSION CR603463
VERSION 1 GI:50484270
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1854)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1854)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1. .1854
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YA14"
/tissue_type="Adult brain"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 91.1%; Score 1673.4; DB 4; Length 1854;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1762; Conservative 0; Mismatches 1; Indels 58; Gaps 2;
Qy 1 GAGGAGTGGAGATGGCGGCGCGGCTCAGGGGCGGGGCGGGAGCCCGGTAGAA 60
Db 75 GAGGAGTGGAGATGGCGGCGCGGCTCAGGGGCGGGGCGGGAGCCCGGTAGAA 134
Qy 61 CCAGAGGGGCTGGCGCGCGGGGTCCTCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTG 120
Db 135 CCAGAGGGGCTGGCGCGCGGGGTCCTCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTG 194
Qy 121 ACCTGGGCGCGGCTACACGAGTTCAGTTCAGTTCGCGAGGGGCGGCTACGGATGGTCA 180
Db 195 ACCTGGGCGCGGCTACACGAGTTCAGTTCAGTTCGCGAGGGGCGGCTACGGATGGTCA 254
Qy 181 GCTCGGCTTATGACCGTTCGCAAGACTCGCTGGCCATCAAGAGATCAGCCCTTCG 240
Db 255 GCTCGGCTTATGACCGTTCGCAAGACTCGCTGGCCATCAAGAGATCAGCCCTTCG 314
Qy 241 AACATCAGACTTACTGCGAGCGACGCTCCGGGAGATCCAGATCTGCTGGCTTCGCC 300
Db 315 AACATCAGACTTACTGCGAGCGACGCTCCGGGAGATCCAGATCTGCTGGCTTCGCC 374
Qy 301 ATGAGATGTCATCGGATCCGAGACATTCCTGGGGGCTCCACCTCGAAGCCATGAGAG 360
Db 375 ATGAGATGTCATCGGATCCGAGACATTCCTGGGGGCTCCACCTCGAAGCCATGAGAG 434
Qy 361 ATGTCTACATTTGTGAGGACCTGATGGAGATGTAACCTGTGTAAGTTCTGTAAGAGCCAG 420
Db 435 ATGTCTACATTTGTGAGGACCTGATGGAGATGTAACCTGTGTAAGTTCTGTAAGAGCCAG 494
Qy 421 AGCTGAGCAATGACCATATCTGCTACTTCTTCTTACCAGATCTCTCGGGGGCTCAAGTACA 480
```

495	Db		AGCTGAGCAATGACCAATATCTGCTACTTCTCTACCAAGATCTCTGCGGGGCTCAAGTACA	554
481	Qy		TCCACTCGGCCAAGTGTCTCCACCGAGATCTAAGAGCCTCCAACTGCTCATCAACACCA	540
555	Db		TCCACTCGGCCAAGTGTCTCCACCGAGATCTAAGAGCCTCCAACTGCTCATCAACACCA	614
541	Qy		CTTGCGACCTTAAGATTTGTGATTTTCGGCTTGCGCCCGGATTTGCCGATCTCTGAGCATGAC	600
615	Db		CTTGCGACCTTAAGATTTGTGATTTTCGGCTTGCGCCCGGATTTGCCGATCTCTGAGCATGAC	674
601	Qy		ACACCGGCTTCTTGACCGAGTATGTGCTTACGGCTTGCGCCCGGATTTGCCGATCTCTGAGCATGAC	660
675	Db		ACACCGGCTTCTTGACCGAGTATGTGCTTACGGCTTGCGCCCGGATTTGCCGATCTCTGAGCATGAC	734
661	Qy		TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTTGCTGTGTGGGCTGCATTTCTGGCTG	720
735	Db		TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTTGCTGTGTGGGCTGCATTTCTGGCTG	794
721	Qy		AGATGCTCTTAACCGGCCCATCTTTCCTTGCGAAGCACTACCTTGGAATCAACACCA	780
795	Db		AGATGCTCTTAACCGGCCCATCTTTCCTTGCGAAGCACTACCTTGGAATCAACACCA	854
781	Qy		TTCTGGGCATCTGGGCTCCCAATCCAGAGGACCTGGAATTTGATCATCAATGAAGG	840
855	Db		TTCTGGGCATCTGGGCTCCCAATCCAGAGGACCTGGAATTTGATCATCAATGAAGG	914
841	Qy		CCCGAACTACTACAGTCTCTGCCCTCCAAAGCAAGTGGCTTTGGGCCAAGCTTTTCC	900
915	Db		CCCGAACTACTACAGTCTCTGCCCTCCAAAGCAAGTGGCTTTGGGCCAAGCTTTTCC	974
901	Qy		CCAACTCAGACTCCAAA-----GCCCTTGACCTGTGTGACCGGATGTT	943
975	Db		CCAACTCAGACTCCAAAGTTTCGAGAGACATGAGGCCCTTGACCTGTGACCGGATGTT	1034
944	Qy		AACCTTTAAACCCCAATAAACGAGTACAGTGGAGGAAGCGTGGCTCAACCCCTACCTGGA	1003
1035	Db		AACCTTTAAACCCCAATAAACGAGTACAGTGGAGGAAGCGTGGCTCAACCCCTACCTGGA	1094
1004	Qy		GCAGTACTATGACCCGACGATGAGCCAGTGGCCGAGAGGCCCTTCACTTCGCGCATGGA	1063
1095	Db		GCAGTACTATGACCCGACGATGAGCCAGTGGCCGAGAGGCCCTTCACTTCGCGCATGGA	1154
1064	Qy		GCTGGATGACCTACTTAAGGAGCGGCTGAAGAGCTCATCTTCCAGAGAGACAGCACGCTT	1123
1155	Db		GCTGGATGACCTACTTAAGGAGCGGCTGAAGAGCTCATCTTCCAGAGAGACAGCACGCTT	1214
1124	Qy		CCAGCCCGGAGTGTGGAGGCCCTTAGCCAGACAGACATCTCTGCAACCTTGGGGCTG	1183
1215	Db		CCAGCCCGGAGTGTGGAGGCCCTTAGCCAGACAGACATCTCTGCAACCTTGGGGCTG	1273
1184	Qy		GAA CAGAACTGGCAAGAGGCAAGAGTCACTGAGGGGCTCTGTACCCAGGACCTGCTG	1243
1274	Db		-----GGACCTGCTGCT-----	1283
1244	Qy		CTTGCTGCCCCCTCTCCGCGCAGACTGTTAGAAATGGAACACTGTGCCAGCCCGGACCT	1303
1284	Db		CTTGCTGCCCCCTCTCCGCGCAGACTGTTAGAAATGGAACACTGTGCCAGCCCGGACCT	1343
1304	Qy		TGGCAGCCAGGCGCGGGTGGAGATGGGCTTGCGCCACTCTCTCTCTTTGTGTAGGCTC	1363
1344	Db		TGGCAGCCAGGCGCGGGTGGAGCA-TGGGCTTGCGCCACTCTCTCTCTTTGTGTAGGCTC	1403
1364	Qy		CAGCTTCAGGAGGCGCAAGGCTTCTCTCTCCCAACCGGCCCTCCCAAGGGGCTCGGGA	1423
1404	Db		CAGCTTCAGGAGGCGCAAGGCTTCTCTCTCCCAACCGGCCCTCCCAAGGGGCTCGGGA	1463
1424	Qy		GCTCAGGTGGCCCAAGTTCATCTCCCGCTGCTGCTGCGGCTTACCTTCCCAAGCG	1483
1464	Db		CCTCAGGTGGCCCAAGTTCATCTCCCGCTGCTGCTGCGGCTTACCTTCCCAAGCG	1523
1484	Qy		TCCCAGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGCTGCTGCCCAACTGCTGAAGGC	1543

Db	1524	TCCAGTCTCTGGAGGTTCTTGGAAATGGAAGGGTTCTGGGCTGCCCCACACCTGCTGAAGGGC	1584
Qy	1544	AGAGGTGGAGGGTGGGGGGCGCTGAGTAGGGAAGCTCAGGGCCATGCGCTGCCCGCCCTCATCT	1603
Db	1584	AGAGGTGGAGGGTGGGGGGCGCTGAGTAGGGAAGCTCAGGGCCATGCGCTGCCCGCCCTCATCT	1643
Qy	1604	CATTCAAAACCCCAACCTAGTTCCTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATC	1663
Db	1644	CATTCAAAACCCCAACCTAGTTCCTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATC	1703
Qy	1664	CCTGAGGAGCCAGCGCGGGCGGAATCCCTCCCTGTCAAAAGCTGTCACTTCGCGTGCCT	1723
Db	1704	CCTGAGGAGCCAGCGCGGGCGGAATCCCTCCCTGTCAAAAGCTGTCACTTCGCGTGCCT	1763
Qy	1724	CGCTGCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTTGGGGCGCTGGAGAGCCCGGCGCCC	1783
Db	1764	CGCTGCTTCTGTGTGTGAGCAGAAAGTGGAGCTTGGGGCGCTGGAGAGCCCGGCGCCC	1823
Qy	1784	CTGCGACCTCCCTGACCCCGTCTAATATATAA	1814
Db	1824	CTGCGACCTCCCTGACCCCGTCTAATATATAA	1854
RESULT 7			
CR621977			
LOCUS			
DEFINITION			
CR621977 1729 bp mRNA linear HTC 21-JUL-2004			
full-length cDNA clone CS0DJ015YL15 of T cells (Jurkat cell line)			
Cot 10-normalized of Homo sapiens (human).			
ACCESSION			
CR621977.1 GI:50502784			
VERSION			
HTC; CNSLUT_CDNA.			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE			
1 (bases 1 to 1729)			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
Full-length cDNA libraries and normalization			
JOURNAL			
Unpublished			
REMARK			
Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue			
2 (bases 1 to 1729)			
Genoscope.			
REFERENCE			
AUTHORS			
Genoscope.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (20-JUL-2004) GENOSCOPE - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
- Web : www.genoscope.cns.fr)			
COMMENT			
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.			
FEATURES			
source			
1. .1729			
Location/Qualifiers			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DJ015YL15"			
/tissue_type="T cells (Jurkat cell line) Cot			
10-normalized"			
/plasmid="pCMVSPORT_6"			
ORIGIN			
Query Match			
90.7%; Score 1666.4; DB 4; Length 1729;			
Best Local Similarity			
97.1%; Pred. No. 0;			
Matches 1729; Conservative			
0; Mismatches			
1; Indels			
51; Gaps			
Qy	36	GGCGGGGGCGGGAGCCCGCTAGAACCGAGGGGGTTCGGCCCGGGGTCCTCCGGGGAGGTG	95
Db	1	GGCGGGGGCGGGAGCCCGCTAGAACCGAGGGGGTTCGGCCCGGGGTCCTCCGGGGAGGTG	60

```
QY 96 GAGATGGTGAAGGGGCGACGGTTTCGACGTGGGCGCGCTTACACGACGTTGCGATCATC 155
DB |||||
QY 61 GAGATGGTGAAGGGGCGACGGTTTCGACGTGGGCGCGCTTACACGACGTTGCGATCATC 120
DB |||||
QY 156 GCGAGGGGCGGTACGGCATGGTTCAGCTCGGCGCTATGACACAGTGGCGAAGACTCCGGTG 215
DB |||||
QY 121 GCGAGGGGCGGTACGGCATGGTTCAGCTCGGCGCTATGACACAGTGGCGAAGACTCCGGTG 180
DB |||||
QY 216 GCCATCAAGAAGATCAGGCCCTTCGAACATCAGACCTTACTGCGCAGCGCAGCTCCGGGAG 275
DB |||||
QY 181 GCCATCAAGAAGATCAGGCCCTTCGAACATCAGACCTTACTGCGCAGCGCAGCTCCGGGAG 240
DB |||||
QY 276 ATCCAGATCTGCTGGCGCTTCGGCGATGAGAAATGCTATCGGCAATCGAGACATCTCGGG 335
DB |||||
QY 241 ATCCAGATCTGCTGGCGCTTCGGCGATGAGAAATGCTATCGGCAATCGAGACATCTCGGG 300
DB |||||
QY 336 GCGTCCACCTCGAAGCCATGAGAGATGCTACATTTGTGCGAGACCTGATGAGAGACTGAC 395
DB |||||
QY 301 GCGTCCACCTCGAAGCCATGAGAGATGCTACATTTGTGCGAGACCTGATGAGAGACTGAC 360
DB |||||
QY 396 CTGTACAAATGTTGTGAAGGCCAGCAGCTGAGCAATGACCATATCTGCTACTTCTCTTAC 455
DB |||||
QY 361 CTGTACAAATGTTGTGAAGGCCAGCAGCTGAGCAATGACCATATCTGCTACTTCTCTTAC 420
DB |||||
QY 456 CAGATCTTCGGGGCCTCAAGTATATCCATCTCGGCCAACTGCTCCACCGAGATCTAAAG 515
DB |||||
QY 421 CAGATCTTCGGGGCCTCAAGTATATCCATCTCGGCCAACTGCTCCACCGAGATCTAAAG 480
DB |||||
QY 516 CCCTCCAACTGCTCATCAACACCACTGCGACCTTAAGATTTGTGATTTTCGGCGCTGGCC 575
DB |||||
QY 481 CCCTCCAACTGCTCATCAACACCACTGCGACCTTAAGATTTGTGATTTTCGGCGCTGGCC 540
DB |||||
QY 576 CGGATTTGCCGATCTGAGCATGACCAACCGGCTTCCTGACCGAGATGTTGGCTACGCGC 635
DB |||||
QY 541 CGGATTTGCCGATCTGAGCATGACCAACCGGCTTCCTGACCGAGATGTTGGCTACGCGC 600
DB |||||
QY 636 TGGTACCGGGCCCCAGAGATCATGCTGAATCTCAAGGGCTATACCAAGTCCATCGACATC 695
DB |||||
QY 601 TGGTACCGGGCCCCAGAGATCATGCTGAATCTCAAGGGCTATACCAAGTCCATCGACATC 660
DB |||||
QY 696 TGGTCTGTGGGCTGCATTCTGGCTGAGATGCTCTTAACCGGCGCCATCTTCCCTGGCAAG 755
DB |||||
QY 661 TGGTCTGTGGGCTGCATTCTGGCTGAGATGCTCTTAACCGGCGCCATCTTCCCTGGCAAG 720
DB |||||
QY 756 CACTACCTGATCAGTCAACCAATCTTGGGATCTCTGGGCTCCCGCATCCAGAGGAGC 815
DB |||||
QY 721 CACTACCTGATCAGTCAACCAATCTTGGGATCTCTGGGCTCCCGCATCCAGAGGAGC 780
DB |||||
QY 816 CTGAATTTGTATCATCAACATGAAGGCCCGAACTTACCTACAGTCTCTGCGCTCCAGACC 875
DB |||||
QY 781 CTGAATTTGTATCATCAACATGAAGGCCCGAACTTACCTACAGTCTCTGCGCTCCAGACC 840
DB |||||
QY 876 AAGTGGCTTGGGCAAGCTTTTCCCAAGTCAAGTCCAAAGCCCTTGACCTGCTGGAC 935
DB |||||
QY 841 AAGTGGCTTGGGCAAGCTTTTCCCAAGTCAAGTCCAAAGCCCTTGACCTGCTGGAC 900
DB |||||
QY 936 CGGATGTTAACTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 995
DB |||||
QY 901 CGGATGTTAACTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 960
DB |||||
QY 996 TACCTGGAGCAGTACTATGACCCGACGATGAGCCAGTGGCGGAGGAGCCCTTACCTTC 1055
DB |||||
QY 961 TACCTGGAGCAGTACTATGACCCGACGATGAGCCAGTGGCGGAGGAGCCCTTACCTTC 1020
DB |||||
QY 1056 GCCATGAGCTGGATGATCTTACCTTAAGAGCGGCTGAAGAGGCTCATCTTCCAGGAGACA 1115
DB |||||
QY 1021 GCCATGAGCTGGATGATCTTACCTTAAGAGCGGCTGAAGAGGCTCATCTTCCAGGAGACA 1080
DB |||||
QY 1116 GCAGCTTCCAGCCGGAGTCTGGAGGCCCTTACGCGCCAGACAGATCTCTGCAACCT 1175
DB |||||
QY 1081 GCAGCTTCCAGCCGGAGTCTGGAGGCCCTTACGCGCCAGACAGATCTCTGCAACCT 1140
DB |||||
QY 1176 GGGGCTTGGAAAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGGCTCTGTCAACCCAGG 1235
```

```
DB |||||
QY 1141 GGGGCT-----GG 1149
DB |||||
QY 1236 AACTGCTCTCTGCTGCCCCCTCTCCCGCCAGACTGTTAGAAAATGACACTGTGCCAGC 1295
DB |||||
QY 1150 AACTGCTCTCTGCTGCCCCCTCTCCCGCCAGACTGTTAGAAAATGACACTGTGCCAGC 1209
DB |||||
QY 1296 CCGGACCTGGCAGACCCAGGCCGGGTGAGCATGGGCTTGGCCACTCTCTCTTTGCT 1355
DB |||||
QY 1210 CCGGACCTGGCAGACCCAGGCCGGGTGAGCATGGGCTTGGCCACTCTCTCTTTGCT 1269
DB |||||
QY 1356 GAGGCTCTCAGCTTTCAGGCGAGCCAAAGGCTTCTCTCCGCCACCCGCCCTCCCGCAGGGG 1415
DB |||||
QY 1270 GAGGCTCTCAGCTTTCAGGCGAGCCAAAGGCTTCTCTCCGCCACCCGCCCTCCCGCAGGGG 1329
DB |||||
QY 1416 CTTGCGGAGCTCAGGTGGCCGCCAGTTTCAATCTCCGCTGCTGCTGCTGGCCCTTACCTT 1475
DB |||||
QY 1330 CTTGCGGAGCTCAGGTGGCCGCCAGTTTCAATCTCCGCTGCTGCTGCTGGCCCTTACCTT 1389
DB |||||
QY 1476 CCGCAGCTCCCACTGCTCTGGCAGTTCTCGAATGGAAGGTTCTGGCTGCCCAACTGC 1535
DB |||||
QY 1390 CCGCAGCTCCCACTGCTCTGGCAGTTCTCGAATGGAAGGTTCTGGCTGCCCAACTGC 1449
DB |||||
QY 1536 TGAAGGGCAGAGTGGAGGGTGGGGGCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCC 1595
DB |||||
QY 1450 TGAAGGGCAGAGTGGAGGGTGGGGGCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCC 1509
DB |||||
QY 1596 CTTCACTCTCATTTCAAAACCCACCTTCTAGTTTCCCTGGAAGAACTTCTTAGTCTCAAGGG 1655
DB |||||
QY 1510 CTTCACTCTCATTTCAAAACCCACCTTCTAGTTTCCCTGGAAGAACTTCTTAGTCTCAAGGG 1569
DB |||||
QY 1656 CTAGCATCTCTGAGGAGCCAGGCCGGCGGAATCCCTCTCCCTGTCAAAGTGTCACTTCG 1715
DB |||||
QY 1570 CTAGCATCTCTGAGGAGCCAGGCCGGCGGAATCCCTCTCCCTGTCAAAGTGTCACTTCG 1629
DB |||||
QY 1716 CGTGGCTCTGCTGCTTCTGCTGCTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCC 1775
DB |||||
QY 1630 CGTGGCTCTGCTGCTTCTGCTGCTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCC 1689
DB |||||
QY 1776 CGCGGCTCTGCGACCTCTGACCCCTCTAATATATAA 1815
DB |||||
QY 1690 CGCGGCTCTGCGACCTCTGACCCGCTAATATATAA 1729
DB |||||

RESULT 8
CR621688
LOCUS full-length cDNA clone CS0D1044VF16 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human) .
ACCESSION CR621688
VERSION CR621688.1 GI:50502495
KEYWORDS HTC; CNSLT cdNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1707)
AUTHORS Li W.B., Gruber C., Jesse J., and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
REFERENCE 2 (bases 1 to 1707)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source Location/Qualifiers
1..1707
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI044YF16"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 89.5%; Score 1644.4; DB 4; Length 1707;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1706; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
QY 46 GGGAGCCCGTAGAACCGAGGGGGTCCGGCCCGGGGTCCCGGGGAGGTGGAGATGGTGA 105
Db 1 GGGAGCCCGTAGAACCGAGGGGGTCCGGCCCGGGGTCCCGGGGAGGTGGAGATGGTGA 60
QY 106 AGGGGAGCCGTTGAGAGTGGGCCCCGGCTACAGCAGTTGCAGTACATCGGGAGGGCG 165
Db 61 AGGGGAGCCGTTGAGAGTGGGCCCCGGCTACAGCAGTTGCAGTACATCGGGAGGGCG 120
QY 166 CGTACGGCATGTGTCAGCTCGGCTATGACACAGTGGCAAGACTCGCGTGGCCATCAAGA 225
Db 121 CGTACGGCATGTGTCAGCTCGGCTATGACACAGTGGCAAGACTCGCGTGGCCATCAAGA 180
QY 226 AGATCAGCCCCCTTCGAACATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCC 285
Db 181 AGATCAGCCCCCTTCGAACATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCC 240
QY 286 TGTGCGCTTCGCGCATGAGATGTGTCATCGGCATCCGAGACATTCGCGGGGTCCACCC 345
Db 241 TGTGCGCTTCGCGCATGAGATGTGTCATCGGCATCCGAGACATTCGCGGGGTCCACCC 300
QY 346 TGGAGCCATGAGAGATGTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAACT 405
Db 301 TGGAGCCATGAGAGATGTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAACT 360
QY 406 TGTGAAAAGCCAGCAGCTGAGCAATGAACATATCTGCTACTTCTCTACAGATCCCTGC 465
Db 361 TGTGAAAAGCCAGCAGCTGAGCAATGAACATATCTGCTACTTCTCTACAGATCCCTGC 420
QY 466 GGGGCTCAAGTACATCCACTCGCCCAAGTGTCTCCACCGAGATCTAAAGCCCTCCAAAC 525
Db 421 GGGGCTCAAGTACATCCACTCGCCCAAGTGTCTCCACCGAGATCTAAAGCCCTCCAAAC 480
QY 526 TGTCTATCAACACACCTGGACCTTAAAGATTTGTGATTTGGGCTGGGCCCGGATTCGG 585
Db 481 TGTCTATCAACACACCTGGACCTTAAAGATTTGTGATTTGGGCTGGGCCCGGATTCGG 540
QY 586 ATCTGAGCATGACCAACACCGGCTTCTGACCGGAGTATGGCTACGCGCTGTGTACCGGG 645
Db 541 ATCTGAGCATGACCAACACCGGCTTCTGACCGGAGTATGGCTACGCGCTGTGTACCGGG 600
QY 646 CCCAGAGATCATGCTGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGTGG 705
Db 601 CCCAGAGATCATGCTGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGTGG 660
QY 706 GCTGCATTCTGGCTGAGATGCTCTCTAAACCGGCCCACTCTTCCCTGGCAAGCACTACTCG 765
Db 661 GCTGCATTCTGGCTGAGATGCTCTCTAAACCGGCCCACTCTTCCCTGGCAAGCACTACTCG 720
QY 766 ATCAGCTCAACCAACATTTCTGGGATCCTGGGCTCCCATCCAGGAGGACCTGAAATTGTA 825
Db 721 ATCAGCTCAACCAACATTTCTGGGATCCTGGGCTCCCATCCAGGAGGACCTGAAATTGTA 780
QY 826 TCATCAACATGAAGGCCGGAACCTACCTACAGTCTCTGCGCTCCAGACCAAGTGGCTT 885
Db 781 TCATCAACATGAAGGCCGGAACCTACCTACAGTCTCTGCGCTCCAGACCAAGTGGCTT 840
QY 886 GGGCCCAAGCTTTTCCCCCAAGTCAGACTCCAAAGCCCTTGACCTGTGGACCGGATGTTAA 945

Db 841 GGGCCAAAGCTTTTCCCAGAGTCAGACTCAAAGCCCTTGACCTGTGACCGGATGTTAA 900
QY 946 CCTTTAACCCCAATAAAGCGATCAGTGGAGGAGCGCTGGCTCAACCCCTACCTGGAGC 1005
Db 901 CCTTTAACCCCAATAAAGCGATCAGTGGAGGAGCGCTGGCTCAACCCCTACCTGGAGC 960
QY 1006 AGTACTATGACCCGACGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCGGCATGGAGC 1065
Db 961 AGTACTATGACCCGACGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCGGCATGGAGC 1020
QY 1066 TGGATGACCTACTAAGGAGCGGCTGAAGGAGCTCATTTCCAGGAGCAGCACGCTTCC 1125
Db 1021 TGGATGACCTACTAAGGAGCGGCTGAAGGAGCTCATTTCCAGGAGCAGCACGCTTCC 1080
QY 1126 AGCCCGGAGTGTGAGGCGCCCTAGCCAGACACATCTCTGACCCCTGGGCGCTGGA 1185
Db 1081 AGCCCGGAGTGTGAGGCGCCCTAGCCAGACACATCTCTGACCCCTGGGCGCT--- 1137
QY 1186 ACAGAACTGGCAAAGGCAAGAGTCACTGAGGGCCTCTGTACCCAGGACCTGCTCC 1245
Db 1138 -----GGACCTGCTCC 1149
QY 1246 TGCCTGCCCTCTCCCGCAGACTGTTAGAAATGGAACACTGTGCCCAGACCCGGAACCTTG 1305
Db 1150 TGCCTGCCCTCTCCCGCAGACTGTTAGAAATGGAACACTGTGCCCAGACCCGGAACCTTG 1209
QY 1306 GAGCCGAGCGCGGGTGGAGCATGGGCTGGCCACCTCTCTCTTGTGCTGAGGCGCTCCA 1365
Db 1210 GCAGCCGAGCGCGGGTGGAGCATGGGCTGGCCACCTCTCTCTTGTGCTGAGGCGCTCCA 1269
QY 1366 GCTTCAGGCGAGCCAAAGCGCTTCTCTCCACCCGCGCTCCCGCAGCGGGCTCGGGAGC 1425
Db 1270 GCTTCAGGCGAGCCAAAGCGCTTCTCTCCACCCGCGCTCCCGCAGCGGGCTCGGGAGC 1329
QY 1426 TCAGTGGCGCCCAAGTTCAATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1485
Db 1330 TCAGTGGCGCCCAAGTTCAATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1389
QY 1486 CCAGTCTCTGGCGAGTTCTGGATGGAAGGTTCTGGCTGGCCCAACCTGCTGAAGGGCAG 1545
Db 1390 CCAGTCTCTGGCGAGTTCTGGATGGAAGGTTCTGGCTGGCCCAACCTGCTGAAGGGCAG 1449
QY 1546 AGGTGAGGGTGGGGGGCTGAGTAGGAGCTCAGGGGCAATGCTGCGCCCTCATCTCA 1605
Db 1450 AGGTGAGGGTGGGGGGCTGAGTAGGAGCTCAGGGGCAATGCTGCGCCCTCATCTCA 1509
QY 1606 TTCAAACCCCAACCTTAGTTTCCCTGAAGAAACATTCCTTAGTCTCAAGGGCTAGCATCC 1665
Db 1510 TTCAAACCCCAACCTTAGTTTCCCTGAAGAAACATTCCTTAGTCTCAAGGGCTAGCATCC 1569
QY 1666 TGAGGAGCCAGCGCGGCGGATCCCTCCCTGTCAAAAGCTGCTCACTTCGCGTGGCCCTCG 1725
Db 1570 TGAGGAGCCAGCGCGGCGGATCCCTCCCTGTCAAAAGCTGCTCACTTCGCGTGGCCCTCG 1629
QY 1726 CTGCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGCGCCCT 1785
Db 1630 CTGCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGCGCCCT 1689
QY 1786 GCCACCTCCCTGACCCGT 1803
Db 1690 GCCACCTCCCTGACCCGT 1707

RESULT 9
CR600245
LOCUS
DEFINITION
CR600245
ACCESSION
VERSION
KEYWORDS
SOURCE
CR600245
full-length cDNA clone CS0C016YM24 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
CR600245
CR600245.1 GI:50481052
HTC; CNSLT_cDNA.
Homo sapiens (human)

ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1 (bases 1 to 1678)									
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.									
TITLE	Full-length cDNA libraries and normalization									
JOURNAL	Unpublished									
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue									
REFERENCE	2 (bases 1 to 1678)									
AUTHORS	Genoscope.									
TITLE	Direct Submission									
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)									
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.									
FEATURES	Location/Qualifiers									
source	1..1678									
	/organism="Homo sapiens"									
	/mol_type="mRNA"									
	/db_xref="taxon:9606"									
	/clone="CS0DC016YM24"									
	/tissue_type="Neuroblastoma Cot 25-normalized"									
	/plasmid="pCMVSPORT_6"									
ORIGIN										
	Query Match	87.9%;	Score	1615.4;	DB 4;	Length	1678;			
	Best Local Similarity	97.0%;	Pred. No.	0;						
	Matches	1677;	Conservative	0;	Mismatches	1;	Indels	51;	Gaps	1;
Qy	73	GCCCGGGGTCCCGGGGAGGTGAGATGTGTAAGGGGCGACCGTTTCGAGTGGGCCCGC	132							
Db	1	GCCCGGGGTCCCGGGGAGGTGAGATGTGTAAGGGGCGACCGTTTCGAGTGGGCCCGC	60							
Qy	133	GCTACACGAGTTCAGTACATCGCGGAGGCGGTACGGCATGTCAGCTCGGCCCTATG	192							
Db	61	GCTACACGAGTTCAGTACATCGCGGAGGCGGTACGGCATGTCAGCTCGGCCCTATG	120							
Qy	193	ACCACTGCGCAAGACTCGCTGGGCCCATCAAGAAGATCAGCCCCCTTCGAACATCAGACCT	252							
Db	121	ACCACTGCGCAAGACTCGCTGGGCCCATCAAGAAGATCAGCCCCCTTCGAACATCAGACCT	180							
Qy	253	ACTGCCAGCGCACGCTCGGGAGATCCAGATCTCTGCTGCGCTTCGGCCATGAGAAATGCA	312							
Db	181	ACTGCCAGCGCACGCTCGGGAGATCCAGATCTCTGCTGCGCTTCGGCCATGAGAAATGCA	240							
Qy	313	TCGGCATCCGAGACTTCGGGGCGTCCACCTCGGAAGCCATCAGAGATGCTACATATG	372							
Db	241	TCGGCATCCGAGACTTCGGGGCGTCCACCTCGGAAGCCATCAGAGATGCTACATATG	300							
Qy	373	TGCAGGACTCATGAGAGACTGACCTGTACAAGTTGCTGAAAGCCAGCAGCTGAGCAATG	432							
Db	301	TGCAGGACTCATGAGAGACTGACCTGTACAAGTTGCTGAAAGCCAGCAGCTGAGCAATG	360							
Qy	433	ACCATATCTGTACTTCTCTACAGATCTCTGGGGGGCTCAAGTACATTCACCTCCGCCA	492							
Db	361	ACCATATCTGTACTTCTCTACAGATCTCTGGGGGGCTCAAGTACATTCACCTCCGCCA	420							
Qy	493	ACGTGCTCCACGGAGATCTAAAGCCCTCCACCTGCTCATCAACACCACTCGCACCTTA	552							
Db	421	ACGTGCTCCACGGAGATCTAAAGCCCTCCACCTGCTCATCAACACCACTCGCACCTTA	480							
Qy	553	AGATTGTGATTTTGGGCTTGGCCGATTTGCCGATCTCTTGAGCATGACACACCGGCTTCC	612							
Db	481	AGATTGTGATTTTGGGCTTGGCCGATTTGCCGATCTCTTGAGCATGACACACCGGCTTCC	540							
Qy	613	TGACGGAGTATGTGGCTACGCGCTGGTACCGGGGCCCCAGAGATCATGCTGAACTCCGAGG	672							

Db	1570	TCCTGTCAAAAGCTGTCACTTCGGCGTGCCTTCGTGCTTCTGTGTGTGGTGAGCAGAAGT	1629
Qy	1753	GGAGCTGGGGGGCGGTGGAGAGCCCGCGGCCCTCCACCTCCCTCGACCC	1801
Db	1630	GGAGCTGGGGGGCGGTGGAGAGCCCGCGGCCCTGCCACCTCCCTGACCC	1678
RESULT 10			
CR596064		1718 bp mRNA linear HTC 21-JUL-2004	
LOCUS		full-length cDNA clone CSODL009YH16 of B cells (Ramos cell line)	
DEFINITION		Cot 25-normalized of Homo sapiens (human).	
ACCESSION		CR596064	
VERSION		CR596064.1 GI:50476871	
KEYWORDS		HTC; CNSLT cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE		Full-length cDNA libraries and normalization	
JOURNAL		Unpublished	
REMARK		Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue	
REFERENCE		2 (bases 1 to 1718)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (20-JUL-2004) Genoscope - Centre National de Sequences : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
COMMENT		1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	
FEATURES		Location/Qualifiers	
source		1..1718	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODL009YH16"	
		/tissue type="B cells (Ramos cell line) Cot 25-normalized"	
		/plasmid="pCMVSPORT_6"	
ORIGIN			
Query Match		85.9%; Score 1577.4; DB 4; Length 1718;	
Best Local Similarity		93.8%; Pred. No. 0;	
Matches 1710; Conservative		0; Mismatches 1; Indels 112; Gaps 2;	
Qy	1	GAGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGCGGGAGCCCGTAGAA	60
Db	8	GAGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGCGGGAGCCCGTAGAA	67
Qy	61	CCGAGGGGGTCTCGGCCCGGGGTCCCGGGGAGGTGGAGATGCTCAAGGGGAGCGCGTTCG	120
Db	68	CCGAGGGGGTCTCGGCCCGGGGTCCCGGGGAGGTGGAGATGCTGAAGGGGAGCGTTCG	127
Qy	121	ACGTGGCGCCGCGCTACACGCAGTTGCATGATCATCGGGAGGGGCGCGTAGCGCATGTGTCA	180
Db	128	ACGTGGCGCCGCGCTACACGCAGTTGCATGATCATCGGGAGGGGCGGTACGCGCATGTGTCA	187
Qy	181	GCTCGGGCTATGACACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCG	240
Db	188	GCTCGGGCTATGACACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCG	247
Qy	241	AACATCAGACCTACTTGCAGCGCACGCTCGGGAGATCCAGATCCTGTGCGCTTCGCCC	300
Db	248	AACATCAGACCTACTTGCAGCGCACGCTCGGGAGATCCAGATCCTGTGCGCTTCGCCC	307
Qy	301	ATGAGATGTTCATCGGATCCGAGACATTTGCGGGGGTCCACCTTGGAGACATGAGAG	360

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:30:54 ; Search time 391.141 Seconds
(without alignments)
8348.349 Million cell updates/sec

Title: US-10-623-108-5

Perfect score: 1837

Sequence: 1 gagagtgagatggcgccg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PPUS_COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1708.4	93.0	1866	3	US-08-909-742-1
2	1708.4	93.0	1866	3	US-09-412-289-1
3	1138.4	62.0	1896	3	US-09-417-197-38
4	1025	55.8	1747	2	US-08-176-620A-1
5	1025	55.8	1747	2	US-08-463-862-1
6	1025	55.8	1747	2	US-08-461-985-1
7	1025	55.8	1747	2	US-08-458-887-1
8	1025	55.8	1747	3	US-08-932-787B-1
9	1025	55.8	1747	3	US-08-932-012C-1
10	1025	55.8	1747	3	US-08-888-818C-1
11	958.6	52.2	1875	3	US-09-417-197-56
12	651.4	35.5	1701	3	US-09-949-016-4048
13	646.6	35.2	1611	3	US-08-909-742-2
14	646.6	35.2	1611	3	US-09-412-289-2
15	642	34.9	2791	3	US-09-016-434-1263
16	632.4	34.4	1467	2	US-08-176-620A-3
17	632.4	34.4	1467	2	US-08-463-862-3
18	632.4	34.4	1467	2	US-08-461-985-3
19	632.4	34.4	1467	2	US-08-458-887-3
20	632.4	34.4	1467	3	US-08-932-787B-3
21	632.4	34.4	1467	3	US-08-932-012C-3
22	632.4	34.4	1467	3	US-08-888-818C-3
23	632.4	34.4	1815	3	US-09-417-197-58
24	632.4	34.4	1818	3	US-09-417-197-40

c	25	525.8	28.6	546	3	US-09-854-133-457	Sequence 457, App
	26	494	26.9	574	3	US-09-854-133-456	Sequence 456, App
	27	454.6	24.7	549	2	US-08-176-620A-7	Sequence 7, Appli
	28	454.6	24.7	549	2	US-08-461-985-7	Sequence 7, Appli
	29	454.6	24.7	549	3	US-08-932-787B-15	Sequence 15, Appl
	30	454.6	24.7	549	3	US-08-932-012C-15	Sequence 15, Appl
	31	454.6	24.7	549	3	US-08-888-818C-15	Sequence 15, Appl
	32	383.4	20.9	537	2	US-08-176-620A-9	Sequence 9, Appli
	33	383.4	20.9	537	2	US-08-461-985-9	Sequence 9, Appli
	34	383.4	20.9	537	3	US-08-932-787B-17	Sequence 17, Appl
	35	383.4	20.9	537	3	US-08-932-012C-17	Sequence 17, Appl
	36	383.4	20.9	537	3	US-08-888-818C-17	Sequence 3926, Ap
	37	320.4	17.4	2826	3	US-09-949-016-3926	Sequence 1458, Ap
	38	320.4	17.4	2828	3	US-09-016-434-1458	Sequence 11, Appl
	39	320.4	17.4	2980	3	US-09-266-225D-11	Sequence 1, Appli
	40	312.4	17.0	1502	2	US-08-651-940-1	Sequence 1, Appli
	41	312.4	17.0	1502	3	US-09-295-029-1	Sequence 1, Appli
	42	312.4	17.0	1502	3	US-09-724-768-1	Sequence 1, Appli
	43	307.6	16.7	1334	3	US-09-949-016-3997	Sequence 3997, Ap
	44	307.6	16.7	1310	3	US-09-047-288-1	Sequence 1, Appli
	45	307.6	16.7	1310	3	US-08-802-191-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-909-742-1

; Sequence 1, Application US/08909742

; Patent No. 6007991

; GENERAL INFORMATION:

; APPLICANT: Vimala S. Sivaraman

; APPLICANT: Hsien-Yu Wang

; APPLICANT: Craig C. Malbon

; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES FOR MITOGEN-

; TITLE OF INVENTION: ACTIVATED PROTEIN KINASES AS THERAPY FOR

; TITLE OF INVENTION: BREAST CANCER

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann & Barton, LLP

; STREET: 350 Jericho Turnpike

; CITY: Jericho

; STATE: New York

; COUNTRY: USA

; ZIP: 11753

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for windows

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/909,742

; FILING DATE: August 12, 1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/831,994

; FILING DATE: April 1, 1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/827,520

; FILING DATE: March 28, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Adams, Lindsay S.

; REGISTRATION NUMBER: 36,425

; REFERENCE/DOCKET NUMBER: 178-225 CIP II

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550

; TELEFAX: (516) 822-3582

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1866 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear


```

; CURRENT FILING DATE: 1999-10-05
; EARLIER APPLICATION NUMBER: 08/909,742
; EARLIER FILING DATE: 1997-08-12
; EARLIER APPLICATION NUMBER: 08/831,994
; EARLIER FILING DATE: 1997-04-01
; EARLIER APPLICATION NUMBER: 08/827,520
; EARLIER FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1866
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Description of Art
US-09-412-289-1

```

Query Match	93.0%;	Score 1708.4;	DB 3;	Length 1866;
Best Local Similarity	97.0%;	Pred. No. 0;		
Matches 1783; Conservative	0;	Mismatches 1;	Indels 54;	Gaps 2;
Qy	3	GGAGTGGAGATGGCGGGCGGGCTCAGGGGGGCGGGGGCGGGAGCCCGGTAGAACC	62	
Db				
Qy	64	GGAGTGGAGATGGCGGGCGGGCGGGCTCAGGGGGGCGGGGGCGGGAGCCCGGTAGAACC	123	
Db				
Qy	63	GAGGGGTTCGGCCCGGGGGTCCGGGGGAGGTGGAGATGGTGAAGGGGACGCCGCTTCAC	122	
Db				
Qy	124	GAGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGCCGCTTCAC	183	
Db				
Qy	123	GTGGGCGCGCTACACGCAGTTCGAGTACATCGCGGAGGGGGCGGTACGGCATGGTCAGC	182	
Db				
Qy	184	GTGGGCGCGCTACACGCAGTTCGAGTACATCGCGGAGGGGGCGGTACGGCATGGTCAGC	243	
Db				
Qy	183	TCGGCCCTATGACACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGGCCCTTCGAA	242	
Db				
Qy	244	TCGGCCCTATGACACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGGCCCTTCGAA	303	
Db				
Qy	243	CATCAGACCTACTGTCGAGCGCACGCTCCGGGAGATCCAGATCTCTGCTGGCTTCGGCCAT	302	
Db				
Qy	304	CATCAGACCTACTGTCGAGCGCACGCTCCGGGAGATCCAGATCTCTGCTGGCTTCGGCCAT	363	
Db				
Qy	303	GAGAACTGTCATCGGCATCCGAGACATCTCGGGGGTCCACCCTGGGAAGCCATGAGAGAT	362	
Db				
Qy	364	GAGAACTGTCATCGGCATCCGAGACATCTCGGGGGTCCACCCTGGGAAGCCATGAGAGAT	423	
Db				
Qy	363	GTCTACATTTGTCAGGACCTCATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG	422	
Db				
Qy	424	GTCTACATTTGTCAGGACCTCATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG	483	
Db				
Qy	423	CTGAGCAATGACATATCTGCTATCTTCTCTACAGAGTCTCGGGGGGCTCAAGTACATC	482	
Db				
Qy	484	CTGAGCAATGACATATCTGCTATCTTCTCTACAGAGTCTCGGGGGGCTCAAGTACATC	543	
Db				
Qy	483	CACATCCGCAAGTGTCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACC	542	
Db				
Qy	544	CACATCCGCAAGTGTCTCCACCGAGATCTAAAGCCCTCCAACTGCTCAGCNAACACCACC	603	
Db				
Qy	543	TGCGACCTTAAGATTTGTGATTTTGGCTTGGCCCGGATGGCGATCTCCTGAGCATGACCAC	602	
Db				
Qy	604	TGCGACCTTAAGATTTGTGATTTTGGCTTGGCCCGGATGGCGATCTCCTGAGCATGACCAC	663	
Db				
Qy	603	ACCGGCTTCTGA CGGAGTATGTGGCTACCGGCTGGTACCGGGGCCCGAGATCATGCTG	662	
Db				
Qy	664	ACCGGCTTCTGA CGGAGTATGTGGCTACCGGCTGGTACCGGGGCCCGAGATCATGCTG	723	
Db				
Qy	663	AACATCCAAGGGCTATACCAGATGCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTCAG	722	
Db				
Qy	724	AACATCCAAGGGCTATACCAGATGCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTCAG	783	
Db				
Qy	723	ATGCTCTCTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCACTT	782	
Db				
Qy	784	ATGCTCTCTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCACTT	843	
Db				

RESULT 3									
US-09-417-197-38									
; Sequence 38, Application US/09417197									
; Patent No. 6518021									
; GENERAL INFORMATION:									
; APPLICANT: Ole THASTRUP, et al.									
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I									
; FILE REFERENCE: 3759-0110P									
; CURRENT APPLICATION NUMBER: US/09/417,197									
; FILING DATE: 1999-10-07									
; NUMBER OF SEQ ID NOS: 143									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 38									
; LENGTH: 1896									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: EGFP-Erk1 fusion construct									
; NAME/KEY: CDS									
; LOCATION: (1)..(1893)									
US-09-417-197-38									
Query Match									
Best Local Similarity 62.0%; Score 1138.4; DB 3; Length 1896;									
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	12	ATGCGCGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCGTAGAACCAGGGGGGTC	71						
DB	757	ATGCGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCGTAGAACCAGGGGGGTC	816						
QY	72	GGCCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGCTTCACGTGGGCGCG	131						
DB	817	GGCCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGCTTCACGTGGGCGCG	876						
QY	132	CGCTACACGCGAGTTGAGTACATCGCGAGGGCGGTACGGATGTCAGCTCGGCGTAT	191						
DB	877	CGCTACACGCGAGTTGAGTACATCGCGAGGGCGGTACGGATGTCAGCTCGGCGTAT	936						
QY	192	GACCACGTGCGGAGACTCGCGTGGCCATCAAGAAGATCAGGCCCTTCGAAACATCAGACC	251						
DB	937	GACCACGTGCGGAGACTCGCGTGGCCATCAAGAAGATCAGGCCCTTCGAAACATCAGACC	996						
QY	252	TACTGCGGCGAGCTCCGGGAGATCCAGATCCTGCTGGCTTCGCGCATGAGATGTC	311						
DB	997	TACTGCGGCGAGCTCCGGGAGATCCAGATCCTGCTGGCTTCGCGCATGAGATGTC	1056						
QY	312	ATCGGCATCCGAGACATTCCTGGGGGCTCCACCCTGGAAGCCATGAGAGATGTCTACATT	371						
DB	1057	ATCGGCATCCGAGACATTCCTGGGGGCTCCACCCTGGAAGCCATGAGAGATGTCTACATT	1116						
QY	372	GTGCGGACCTGATGGAGACTGACCTGTGTACAGTTGCTGAAAAGCCAGCAGCTGAGCAAT	431						
DB	1117	GTGCGGACCTGATGGAGACTGACCTGTGTACAGTTGCTGAAAAGCCAGCAGCTGAGCAAT	1176						
QY	432	GACCATATCTGCTACTTCTTACAGATCTCGGGGGCGCTCAAGTACATCCATCCGCGC	491						
DB	1177	GACCATATCTGCTACTTCTTACAGATCTCGGGGGCGCTCAAGTACATCCATCCGCGC	1236						
QY	492	AACGTGCTCCAGCAGATCTAAGCCCTTCAACCTGCTCATCAACACCACTGCGACCTT	551						
DB	1237	AACGTGCTCCAGCAGATCTAAGCCCTTCAACCTGCTCATCAACACCACTGCGACCTT	1296						
QY	552	AAGATTGTGATTTCGGCTCGGCCCGGATTCGCGATTCCTGAGCATGACCAACCGGCTTC	611						
DB	1297	AAGATTGTGATTTCGGCTCGGCCCGGATTCGCGATTCCTGAGCATGACCAACCGGCTTC	1356						
QY	612	CTGACCGAGTATGTGGCTACGCGCTGGTACCGGGCCCCAGAGATCATGTGAATCCAG	671						
DB	1357	CTGACCGAGTATGTGGCTACGCGCTGGTACCGGGCCCCAGAGATCATGTGAATCCAG	1416						
QY	672	GGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCTTCTGGCTGAGATGCTCTCT	731						

RESULT 4

US-08-176-620A-1
; Sequence 1, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,620A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: unknown		Query Match		55.8%; Score 1025; DB 2; Length 1747;	
MOLECULE TYPE: DNA (genomic)		Best Local Similarity		78.1%; Pred. No. 2.1e-202;	
FEATURE:		Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;			
NAME/KEY: CDS					
LOCATION: 1..1095					
US-08-176-620A-1					
Qy	58	GAACCGAGGGGTCGGCCCGGGGTCGGGGGAGGTGGAGATGGTGAAGGGGACGGT	117		
Db	5	GAATGCTGGGGTCTCCCGGTGGTCCCGGGAGGTGGAGTGGTGAAGGGGACGGAT	64		
Qy	118	TCGACGTGGGCGCGCTACACGAGTTGACAGTATCGGCGAGGGCGCTACGGGATGG	177		
Db	65	TCGACGTGGGCGCGCTACACGAGTTGACAGTATCGGCGAGGGCGCTACGGGATGG	124		
Qy	178	TCAGCTCGGCTATGACACGTTGGCGGAGAGTCCGCTGGCCATCAAGAAGATCAGCCCT	237		
Db	125	TCAGCTCAGCATATGACACGTTGGCGGAGAGTCCGCTGGCCATCAAGAAGATCAGCCCT	184		
Qy	238	TCGAATCATGACCTACTGCGAGGCGAGCTCCGGGAGATCCAGATCTCTGGCTTCC	297		
Db	185	TCGACGATCAAACTACTGTGCGGCGAGCTCCGGGAGATCCAGATCTCTGGATTC	244		
Qy	298	GCCATGAGATGTCATCGGCGATCCGAGACATTCCTCGGGCGTCCACCCTCGAAGCCATGA	357		
Db	245	GCCATGAGATGTCATAGGAGATCCGAGACATTCCTCGAGACACCCCTCGAAGCCATGA	304		
Qy	358	GAGATGTCTACATGTGCGAGGACCTGATGGAGAGTGAACCTGTATCAAGTTGCTGAAAGCC	417		
Db	305	GAGATGTCTACATGTGCGAGGACCTGATGGAGAGTGAACCTGTATCAAGTTGCTGAAAGCC	364		
Qy	418	AGCAGCTGAGCAATGACATATCTGTCTTCTCTACAGATCTCGGGGGCTCAAGT	477		
Db	365	AGCAGCTGAGCAATGACATATCTGTCTTCTCTACAGATCTCGGGGGCTCAAGT	424		
Qy	478	ACATCGACTCGGCAAGTGTCTCAGCGAGATCTAAGCCCTCCAACTGCTCATCAACA	537		
Db	425	ACATCGACTCGGCAAGTGTCTCAGCGAGATCTAAGCCCTCCAACTGCTCATCAACA	484		
Qy	538	CCACTCGGACTTAAAGTTGATTTGGGCTCGGCGGCGGATTCGCGATCTGAGCATG	597		
Db	485	CCACTCGGACTTAAAGTTGATTTGGGCTCGGCGGCGGATTCGCGATCTGAGCATG	544		
Qy	598	ACCACACCGGCTTCTGACCGAGTATGGGCTACCGGCTGGTACCGGGCCCGAGATCA	657		
Db	545	ACCACACCGGCTTCTGACCGAGTATGGGCTACCGGCTGGTACCGGGCCCGAGATCA	604		
Qy	658	TGCTGAATCCAGGGCTATACCAAGTCCATCGACATCTGTGTGGGCTGGATTCGG	717		
Db	605	TGCTGAATCCAGGGCTATACCAAGTCCATCGACATCTGTGTGGGCTGGATTCGG	664		
Qy	718	CTGAGATGCTCTTAACCGGCGCTATCTCCCTGGCAAGCACTACTGGATCAGCTCAACC	777		
Db	665	CTGAGATGCTCTTAACCGGCGCTATCTCCCTGGCAAGCACTACTGGATCAGCTCAACC	724		
Qy	778	ACATTTCTGGGCTATCTGGGCTATCTCCCAAGGAGACCTGAATTTGATCATCAACATGA	837		
Db	725	ACATTTCTGGGCTATCTGGGCTATCTCCCAAGGAGACCTGAATTTGATCATCAACATGA	784		
Qy	838	AGGCGGAACTACTACAGTCTCTGCGCTCCAGACCAAGGCTGGGCTGGCCAGCTTT	897		
Db	785	AGGCGGAACTACTACAGTCTCTGCGCTCCAGACCAAGGCTGGGCTGGCCAGCTTT	844		
Qy	898	TCCCAAGTCTGACTTCCAAAGCTTTGACCTGTGGACCGGATGTTAACTTTAACCCCA	957		
Db	845	TCCCAAGTCTGACTTCCAAAGCTTTGACCTGTGGACCGGATGTTAACTTTAACCCCA	904		
Qy	958	ATAAACGATCAGAGTGGAGGAGCGGTGGCTCAACCCCTACTCGGAGCAGTATGACC	1017		

RESULT 5

US-08-463-862-1
; Sequence 1, Application US/08463862
; Patent No. 5776751
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G. et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

Db	905	ACAGCGCATCAGTAGAGGAAGCACTGGCTCACCTTTACCTTGGAAACAGTACTATGATC	964
Qy	1018	CGACGGATGAGCGAGTGGCGAGAGCCCTTTCACCTTGGCCATGGAGCTGGATGACCTAC	1077
Db	965	CGACGATGAACCAAGTGGCTGAGGAGCCATTCACCTTTGACATGGAGCTGGATGATCC	1024
Qy	1078	CTAAGGAGCGGCTGAAGGAGCTCATCTCCAGAGACAGCAGCTTCCAGCCCGGAGTGC	1137
Db	1025	CCAAGGAGCGGCTGAAGGAGCTGATCTTCAAGAGACAGCCCTTCCAGCCAGGGGCAC	1084
Qy	1138	TGAGGCGCCCTAGCCCGACAGACATCTCTGACCCCTGGGGCTCGAAACAGAACTGGCA	1197
Db	1085	CAGAGGCCCTTCAACAGAACACACACCTCTGCTTTTGGACCTGG	1131
Qy	1198	AAGAGCAAGAGGTCACTGAGGCGCTCTGTACCCAGGACCTGCTCTGCTGCTGCTCCTC	1257
Db	1132	-----TCTGCTCTACCTGCTCCT	1149
Qy	1258	TCCCGCCAGCTGTTAGAAATGGACACTGTGCGCCAGCCCGGACCTTGGCAGCCAGGCC	1317
Db	1150	TCTCTGCAATGTTTAAAGAAATG--AACTTTGTCTCAACCCGAGCCCGGCGAGCCAGG-C	1206
Qy	1318	GGGCTGGAGCATGGGCTGCGCCACCTCTCTCTTTGCTGAGGCTCCAGCTTCAGGCAGG	1377
Db	1207	TGACCAAGGGTGGGCTGGCACCTCTCTCACTCTGCTGGGCTCTCTCTGTTCAAG--AG	1264
Qy	1378	CCAAGGCTTCTCTCCCAACCGCCCTCCCAAGGGGCTCCGGAGCTCAGTGGGCCCC	1437
Db	1265	GCTTCTCCACCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1324
Qy	1438	AGTTCAATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1497
Db	1325	AGAGCTGATCTCTGCTGCTG--TGTCTTTATCTATCTGCTAGCCCGAGCTCTGTTAGAC	1383
Qy	1498	AGTTCTGGAAATGAAGGTTCTGGCTGCGCCCAACCTGTGAAGGGCAGAGGTGGAGGGTG	1557
Db	1384	GGTTCTGGAAATGAAGGCTATGACCGCCCTA-----GGACCTGTGCTACAGAGGGTG	1437
Qy	1558	GGGGGCGCTGAGTAGGAGCTCAGGGGCTGCTGCGCCCTCATCTCATATTCATCAACCCAC	1617
Db	1438	GAGGCGACTGAGT-----AGGCTAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1489
Qy	1618	CTAGTTTCTGAAAGAACTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGG	1677
Db	1490	CCCATTTCTGACAGAACTTCTTAAATCTCAAGGGCTAGTTTCCCTGAGGAGCCAGC	1549
Qy	1678	CCGGGCGGATCCCTCCCTGTCAAAGCTGTCACTTCCGCTGCTGCTGCTGCTGCTGCTGCT	1737
Db	1550	CTAGGCTTAACTCTCCCTCTC--AAGCTGCCACATGTAACGCTTCTGCTGCTGCTGCT	1608
Qy	1738	GT--GGTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCCGGGCGCCCTGCGACCTCCCT	1796
Db	1609	GTGGTGATTTGGATGTGGAGGCGGGCGCTGGAGAGCCCGTGGCCCTCCCGACCTCCCT	1668
Qy	1797	GACCC---GTCTAATATATAATATAGAGATGTGTCTATGG	1834
Db	1669	GTGCTGTATCTAATATATAATATAGAGATGTGTATATGG	1709

```

/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/463,862
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 435
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 07/701,544
/ FILING DATE: 16-MAY-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mierock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 6526-049
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212 790-9090
/ TELEFAX: 212 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1747 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..1095
/ US-08-463-862-1

```

Query Match 55.8%; Score 1025; DB 2; Length 1747;
 Best Local Similarity 78.1%; Pred. No. 2.1e-202;
 Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

```

QY 58 GAACGAGGGGGTCCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGT 117
DB 5 GAACGCTGGGTCGTCGCGGTGGTCCCGGGGAGGTGGAGGTGGTGAAGGGGCGAGCCAT 64

QY 118 TCAGCTGGGCGCGCGCTACACGAGTTCAGTACATCCGCGAGGCGCGGTACGGCATGG 177
DB 65 TCAGCTGGGCGCGCGCTACACGAGTTCAGTACATCCGCGAGGCGCGGTACGGCATGG 124

QY 178 TCAGCTGGGCGCGCTACACGAGTTCAGTACATCCGCGAGGCGCGGTACGGCATGG 237
DB 125 TCAGCTGGGCGCGCTACACGAGTTCAGTACATCCGCGAGGCGCGGTACGGCATGG 184

QY 238 TCAGCTGGGCGCGCTACACGAGTTCAGTACATCCGCGAGGCGCGGTACGGCATGG 297
DB 185 TCAGCTGGGCGCGCTACACGAGTTCAGTACATCCGCGAGGCGCGGTACGGCATGG 244

QY 298 GCATGAGAAATGTCATCGGCGATCCGAGACATTCCTGGGGGCGTCCACCTGGAGGCCATGA 357
DB 245 GCATGAGAAATGTCATCGGCGATCCGAGACATTCCTGGAGCACCCCTGGAGGCCATGA 304

QY 358 GAGATGTCATATGTCGAGGAGCTGATGGAGACTGACCTGATCAAGTTGCTGAAAAGCC 417
DB 305 GAGATGTCATATGTCGAGGAGCTGATGGAGACTGACCTGATCAAGTTGCTGAAAAGCC 364

QY 418 AGCAGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTGGGGGCGCTCAAGT 477
DB 365 AGCAGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTGGGGGCGCTCAAGT 424

QY 478 ACATCCACTCCGCAAGGTGCTCCACCGAGATCTAAAGGCCCTCAACCTGCTCATCAACA 537
DB 425 ACATCCACTCCGCAAGGTGCTCCACCGAGATCTAAAGGCCCTCAACCTGCTCATCAACA 484

QY 538 CCACCTCGGACCTTAAGATTGATTTGGCCCTGGCCCGGATGGCGGATCTCGGACATG 597
DB 485 CCACCTCGGACCTTAAGATTGATTTGGCCCTGGCCCGGATGGCGGATCTCGGACATG 544

```

```

QY 598 ACCACACCGGGCTTCTGACGAGGATATGTGGCTACCGCGCTGGTACCGGGGCCCGAGATCA 657
DB 545 ACCACACCTGGCTTCTGACCGAGTATGTGGCCACACGCTGGTACCGAGCCCGAGATCA 604

QY 658 TGCCTGAATCCNAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGATCTCGG 717
DB 605 TGCCTGAATCCNAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGATCTCGG 664

QY 718 CTGAGATGCTCTCTAAACCGGCCCATCTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACC 777
DB 665 CTGAGATGCTCTCTAAACCGGCCCATCTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACC 724

QY 778 ACATTTCTGGGCATCTCTGGGCTCCCATCCAGGAGGACCTGAAATGTATCATCAACATGA 837
DB 725 ACATTTCTAGGTATATCTGGGTTCCCATCCCAAGAGGACCTAAATTTGTATCATTAACATGA 784

QY 838 AGGCCCGAAATCTACCTACAGTCTCTGGCTTCCCAAGCAAGGTGGCTTTGGGCGAAGCTTT 897
DB 785 AGGCCCGAAATCTACCTACAGTCTCTGGCTTCCCAAGCAAGGTGGCTTTGGGCGAAGCTTT 844

QY 898 TCCCAAGTTCAGACTCCAAAGCCCTTGACCTTGCTGACCGGATGTTAACTTTAAACCCCA 957
DB 845 TTCCCAATCTGACTCCAAAGCTCTTGACCTGCTGGAACGATGTTAACTTTAAACCCCA 904

QY 958 ATAAACGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGGATGTAATGAC 1017
DB 905 ACAAGCGCATCACAGTGGAGGAGCACTGGCTCACCCCTTACCTGGAAACAGTACTATGATC 964

QY 1018 CGACGGATGAGCGAGTGGCGGAGGAGCTTTCACCTTCCCATGAGCTGGATGACCTTAC 1077
DB 965 CGACGATGAACCAAGTGGCTGAGGAGCACTTCCCTTTGACATGAGGCTGGATGATCTCC 1024

QY 1078 CTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGC 1137
DB 1025 CCAAGGAGCGGCTGAAGGAGCTGATCTTCCAGAGAGAGCGCGCTTCCAGCCAGGGGAC 1084

QY 1138 TGGAGCGCCCTTAGCCACAGACAGATCTCTGCAACCTTGGGGCTTGGAAACATGGCA 1197
DB 1085 CAGAGCGCCCTTAAACAGAAACAGACACCGCTGCTCTTTTGGACCTGG----- 1131

QY 1198 AAGAGGCAAGGTCATCTGAGGGGCTCTGTACCCAGGACCTGCTCTGCTGCGCCCTC 1257
DB 1132 -----TCTGCTCTACTCTGCTCTCT 1149

QY 1258 TCCCGCCAGACTGTTAGAAAATGGACACTGTGCCAGCCCGGAGCTTGGCAGCCCGAGGCC 1317
DB 1150 TCTCTGCGATGTTAGAAAATG--AACTTGTCTCAACCCGAGACCCCGGACCCAGG-C 1206

QY 1318 GGGGTGGAGCATGGGCTTGGCCACCTCTCTCTTTGTGAGGCGCTCCAGCTTTCAGGACGG 1377
DB 1207 TGGACCAAGGGTGGGCGCTGGCACCCCTCTCACTCTGCTGGGGTCTCTCGTTCAAG--AG 1264

QY 1378 CCAAGGCTTCTCTCCCGACCGGCTTCCCGACCGGGGCTCGGAGCTCAGGTGGGCCCC 1437
DB 1265 GCTTCTCCCACTCCAGTCCCTGCGCCCATCTCCCTTGA CTTGAGTGAAGGTGGTCTCC 1324

QY 1438 AGTTCAATCTCCCGCTGCTGCTGGCGCTTACTCTTCCCGAGGCTCCCACTCTCTGCG 1497
DB 1325 AGAGCTGATCTCTGCTGCTG-TGTCTTTATCTATCTCTGCTAGAGCCAGCTCTGGTAGAC 1383

QY 1498 AGTTCTGGAATGGAAGGGTTCGGCTGCGCCCAACCTGTCTGAAGGCGAGAGGTGGAGGGTG 1557
DB 1384 GGTTCGGAATGGAAGGGCTATGACCGCCCTA-----GGACCTGTGCTACAGAGGGGTG 1437

QY 1558 GGGGGGCTGAGTAGGAGCTCAGGGGCGATGCTGCGCCCTCATCTCATTTCAAAACCCAC 1617
DB 1438 GAGGGGCTGAGT-----AGGCTAAGCTCTGCGCTACTCATCTGCTGGAACCCAC 1489

QY 1618 CTTAGTTTCCCTGAAGGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTTGAGGAGCCAGG 1677
DB 1490 CCAATTTTCCCTGACAGAACATTTCTTAATCTCAAGGGCTAGTTTCCCTTGAGAGCCAGC 1549

QY 1678 CCGGGCCGAATCCCTCCCTGTCAAAAGCTGTCACTTTCGCGTGCCTCGCTCTGCTGT 1737

```

```

Db      1550 CTAGGCGCTAACCCCTCTCCCTCTC-AAAGTCCCAATGTAAAGCCCTTGCTGCTTCTGTGT 1608
Qy      1738 GT--GGTGAGCAGAGTGAGCTGGGGGGCGTGGAGAGCCCGGGCGCCCTGCGCACCTCCCT 1796
Db      1609 GTGGGTGATGGATGGAGCGGGGGCGCTGGAGAGCCCGTGGAGAGCCCGTCCCTCCCACTCCCT 1668
Qy      1797 GACCC---GTCTAATATATAATATATAGAGATGTGTCTATGG 1834
Db      1669 GTCCCTGTATCTAATATATAATATATAGAGATGTGTATATGG 1709

```

RESULT 6

US-08-461-985-1

; Sequence 1, Application US/08461985

; Patent No. 5872006

; GENERAL INFORMATION:

; APPLICANT: Boulton, Teri G.

; APPLICANT: Cobb, Melanie H.

; APPLICANT: Yancopoulos, George D.

; APPLICANT: Nye, Steven

; APPLICANT: Panayotatos, Nikos

; TITLE OF INVENTION: A Family of Map2 Protein Kinases

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/461,985

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/176,620

; FILING DATE: 03-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 6526-123

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-8864/9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1747 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1095

; US-08-461-985-1

Query Match 55.8%; Score 1025; DB 2; Length 1747;
 Best Local Similarity 78.1%; Pred. No. 2.1e-202;
 Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

```

Qy      58 GAACCGAGGGGGTCCGGCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGT 117
Db      5 GAATCTGGGGTCTCCCGGTGGTCCCGGGGAGGTGGAGGTGGTGAAGGGGCGAGCCAT 64
Qy      118 TCAGCGTGGGCGCCCGCTACACGAGTTGACGATCATCGGCGAGGGCGGTACGGCATGG 177

```

```

Db      65 TCAGCGTGGGCGCCACGCTACACGCGCTGCAGTACATCGGCGAGGGGCGGTACGGCATGG 124
Qy      178 TCAGCTCGGCCCTATATGACCAAGTCGCGCAAGATCTCGCGTGGCCATCAAGAAGATCAGCCCT 237
Db      125 TCAGCTCAGCATATGACCAAGTCGCGCAAGATCAGAGTGGCTATCAAGAAGATCAGCCCT 184
Qy      238 TCGAACATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCTCTGCTGCGCTTC 297
Db      185 TCGAGCATCAAACTTACTGTGCGCGCACGCTGAGAGAAATCCAGATCTTGTCTGGATTC 244
Qy      298 GCCATGAGAATGTCTATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTCGGAAGCCATGA 357
Db      245 GCCATGAGAATGTCTATAGGCATCCGAGACATCTCTCAGAGCAACCCACCTTGAAGCCATGA 304
Qy      358 GAGATGTCTACATTTGTGAGGAGCTGATGAGAGCTGACCTGTACACCTGTACAAAGTGTCTGAAAAGCC 417
Db      305 GAGATGTTTACATTTGTTTCAAGGACCTCATGGAGACGACCTGTACAAAGTGTCTTAAAGACC 364
Qy      418 AGCAGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTGCGGGGCTCAAGT 477
Db      365 AGCAGCTGAGCAATGACCAATCTGCTACTCTCTACAGATCTCTTCCGGGGCTCAAGT 424
Qy      478 ACATCCACTCCGCCAACGTGCTCCACCGAGATCTTAAAGCCCTCCAACTGCTCATCAACA 537
Db      425 ACATACACTCGGCCAATGTGCTGCACCGGACCTGAAGCCCTCAATCTGCTTATCAACA 484
Qy      538 CCACCTGCGACCTTAAGATTTTGGGCTTGGGCTGCGGAGTGGCGATGTCCTGAGCATG 597
Db      485 CCACCTGCGACCTTAAGATTTTGGGCTTGGGCTGCGGAGTGGCGATGTCGACCTGAGCAG 544
Qy      598 ACCACACCGGCTTCTGACGAGATGTGCTGCTACGCGCTGGTACCGGGCCCGCAGAGATCA 657
Db      545 ACCACACTGCGCTTTCTGACCGAGTATGTGGCCACACGCTGGTACCGGGCCCGCAGAGATCA 604
Qy      658 TGCTGAATCCCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTGG 717
Db      605 TGCTTAACTCCCAAGGGCTACACCAATCAATTGACATCTGGTCTGTGGGCTGCAATCTGG 664
Qy      718 CTGAGATGCTCTCTAAACCGGCGCATCTTCCCTGGGCAAGCACTACCTGGATCAGCTCAACC 777
Db      665 CTGAGATGCTCTCTAAACCGGCGCATCTTCCCGGCAAGCACTACCTGGACCACTCAACC 724
Qy      778 ACATTTCTGGGCTCTGGGCTCCCATCCAGAGGAGCACTGATTTCTATCATCAACATGA 837
Db      725 ACATTTCTAGGTATCTAGGTTCCTCCATCCCAAGAGGACCTAAATTTGATCTATTAACATGA 784
Qy      838 AGGCCCGAAACTACCTTACAGTCTCTGCGCTCCAAAGACCAAGGTGGCTTGGGCGCAAGCTTT 897
Db      785 AGGCCCGAAACTACCTTACAGTCTCTGCGCTCTTAAACCAAGGTGGCTTGGGCGCAAGCTTT 844
Qy      898 TCCCGCAAGTCAAGCTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAAACCTTTAACCCCA 957
Db      845 TTCCCAAACTGACTCCAAAGCTCTTTGACCTGCTGGACCGGATGTTAAACCTTTAACCCCA 904
Qy      958 ATAAACGAGATCAAGTGGAGGAGCGCTGCTACCTACCTCTGAGGAGCAGTACTATGACC 1017
Db      905 ACAAGCGCATCAAGTAGAGGAGCACTGGCTCACCTTTACCTCTGGAACAGTACTATGATC 964
Qy      1018 CGACGGATGAGCCAGTGGCGGAGCGCTTACCTTTCACCTTTCGCGATGGAGCTGGATGACCTAC 1077
Db      965 CGACAGATGAACCAAGTGGCTGAGAGGCCATTTACCTTTGACATGGAGCTGGATGATCTCC 1024
Qy      1078 CTAAAGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCACGCTTTCAGCCCGGAGTGC 1137
Db      1025 CCAAGGAGCGGCTGAAGGAGCTGATCTTCCAAAGAGACAGCCCGCTTCCAGCCAGGGGCAC 1084
Qy      1138 TGAGGCGCCCTTAGCCAGACAGACATCTCTGACCTTGGGGGCTTGAACAGACTGGCA 1197
Db      1085 CAGAGGCGCCCTTAAAGAAACAGACACCCCTGTCTTTTGGACCTGG----- 1131
Qy      1198 AAGAGGCAAGAGGTCACTGAGGGGCTCTGTGTACCCAGGAGCTTGCCTCTGCTGCGCCCTC 1257
Db      1132 -----TCTGCTCTACCTGCTCTCT 1149

```


Db 785 AGGCCGGAATCTACAGTCTCTGCCCTCTAAACCAAGGTGGCTTGGGCCCAAGCTTT 844
Qy 898 TCCCAAGTCAGACTCCCAAGCCCTTGACCTGTGGAACCGATGTTAACTTTAAACCCCA 957
Db 845 TTCCCAATCTGACTCCAAAGCTCTTGACCTGTGACCGGATGTTAACTTTAAACCCCA 904
Qy 958 ATAAACGATACAGTGGAGAGCGCTGGCTCAACCTCTACCTGGAGCAGTACTATGACC 1017
Db 905 ACAAGCGCATACAGTAGAGAGCACTGGCTCAACCTCTACCTGGAGCAGTACTATGATC 964
Qy 1018 CGAGGATGACCACTGGCCGAGAGCCCTTCACTTCGTCGTCGTCGTCGTCGTCGTCGTC 1077
Db 965 CGACAGATGAACAGTGGCTGGAGAGCACTGGCTCAACCTCTACCTGGAGCAGTACTATG 1024
Qy 1078 CTAAGGAGCGGCTGAAGAGCTCATCTTCCAGGAGACAGCACCGCTTCCAGCCCGAGTGC 1137
Db 1025 CCAAGGAGCGGCTGAAGAGCTCATCTTCCAGGAGACAGCACCGCTTCCAGCCCGAGG 1084
Qy 1138 TGGAGGCCCTTAGCCAGACAGACATCTCTGCAACCTGGGGCTGGAAAGAACTGGCA 1197
Db 1085 CAGAGGCCCTTAAACAGACAGACACCTCTGCTCTTTGGACCTGG----- 1131
Qy 1198 AAGAGGCAAGAGTCACTAGGGGCTCTGTCAACCCAGGACCTGCTCTGCTGCCCTC 1257
Db 1132 -----TCTGCTCTACCTGCTCT 1149
Qy 1258 TCCCGCCAGACTGTTAGAAATGAGACTGTGCCAGCCGGACCTTGGCAGCCAGGCC 1317
Db 1150 TCTCTGAGATGTTAGAAATG--AACTTTGTCTAAACCCGGACCCCGGAGCCAGG-C 1206
Qy 1318 GGGGTGAGCATGGGCTGGCCACTCTCTCTTTGTGAGGGCTCCAGCTTCAGGCAGG 1377
Db 1207 TGAACCAAGGGTGGGCTTGGCAACCTCTCACTCTGCTGGGCTCTCTGCTCAAG--AG 1264
Qy 1378 CCAAGGCTTCTCTCTCCCAACCCGCTCTCCCAAGGGGCTTCGGGAGCTCAGTGGCCCC 1437
Db 1265 GCTTCTCCCACTCCAGTCCCTGCCCCATCTCCCTTTGAACCTGAGTGAAGTGGTCCC 1324
Qy 1438 AGTTCAATCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1497
Db 1325 AGAGCTGATCTGCTGCTGCTG-TGCTTTTATCTATCTCCCTGCTAGCCCCAGCTCTGCTAGAC 1383
Qy 1498 AGTTCTGAATGGAAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1557
Db 1384 GGTCTGGAATGGAAGGGCTATGACCGCCCTA-----GGACCTGTGCTACAGAGGGGTG 1437
Qy 1558 GGGGGGCTGAGTAGGAGCTCAGGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617
Db 1438 GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCTCTCTCTCTCTCTCTCTCTCTCT 1489
Qy 1618 CCTAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGG 1677
Db 1490 CCCATTTTCCCTGACAGAACATTCCTAAATCTCAAGGGCTAGTTTCCCTGAGGAGCCAGC 1549
Qy 1678 CCGGGCCGAATCCCTCTCCCTGCTCAAGAGTGTCACTCTCGCGTCCCTGCTGCTCTGCTGT 1737
Db 1550 CTAGGCTTAACCTCTCTCCCTCTC--AAGTGGCAATGTAAGCCCTTGTGCTGCTGTGT 1608
Qy 1738 GT--GGTGAAGCAAGTGGAGCTGGGGGGCGGTGGAGAGCCCGGCGCCCTGCGCACTCCCT 1796
Db 1609 GTGGGTGATTGATGTGGAGGGGGGGCGGTGGAGAGCCCGTGGCGCCCTGCCCACTCCCT 1668
Qy 1797 GACCC-----GTTCTAATATATAATATAGAGATGTGTCTATGG 1834
Db 1669 GTGCTGTATCTAATATATAATATAGAGATGTGTATATGG 1709

RESULT 8
US-08-932-787B-1
; Sequence 1, Application US/08932787B
; Patent No. 6277963
; GENERAL INFORMATION:

; APPLICANT: Boulton et al.
; TITLE OF INVENTION: ANTIBODIES DIRECTED TOWARD EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASES
; FILE REFERENCE: REG 430-A-1
; CURRENT APPLICATION NUMBER: US/08/932,787B
; CURRENT FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 08/469,547
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/178,488
; PRIOR FILING DATE: 1994-01-07
; PRIOR APPLICATION NUMBER: 07/701,544
; PRIOR FILING DATE: 1991-05-16
; PRIOR APPLICATION NUMBER: 07/532,004
; PRIOR FILING DATE: 1990-06-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: RAT
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1095)
; OTHER INFORMATION: ERK1 cDNA
; US-08-932-787B-1

Query Match 55.8%; Score 1025; DB 3; Length 1747;
Best Local Similarity 78.1%; Pred. No. 2,1e-202;
Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

Qy 58 GAACCGAGGGGTGCGCCCGGGGTCCTCCGGGGAGGTGGAGATGTTGAGGGGCGAGCCGT 117
Db 5 GAATCTCTGGGGTGTCTCCCGGGTGGTCCCGGGAGGTGGAGTGGTGAAGGGGCGAGCAT 64
Qy 118 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 177
Db 65 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 124
Qy 178 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 237
Db 125 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 184
Qy 238 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 297
Db 185 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 244
Qy 298 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 357
Db 245 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 304
Qy 358 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 417
Db 305 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 364
Qy 418 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 477
Db 365 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 424
Qy 478 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 537
Db 425 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 484
Qy 538 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 597
Db 485 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 544
Qy 598 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 657
Db 545 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 604
Qy 658 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 717
Db 605 TCAGCTGCGGCGCTACACGAGTTCAGCTAGTACATCGGCGAGGGCGCTAGCGGATGG 664

```
QY 718 CTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACC 777
Db 665 CTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGACGAGCTCAACC 724

QY 778 ACATTCCTGGGCATCTCTGGCTCCCCATCCAGAGAGACCTGAATTGTATCATCAACATGA 837
Db 725 ACATTCCTGGATATCTCTGGCTTCCCATCCAGAGAGACCTGAATTGTATCATTAACATGA 784

QY 838 AGCCCGAATACCTACAGTCTCTGCTTCCCTCCAGACCAAGTGGCTTGGGCAAGCTTT 897
Db 785 AGCCCGAATACCTACAGTCTCTGCTTCCCTCCAGACCAAGTGGCTTGGGCAAGCTTT 844

QY 898 TCCCCAAGTCAGACTCCAAAGCCCTTGACTCTGTGAGCCGAGTGTAACTTTAAACCCCA 957
Db 845 TTCCCCAATCTGACTCCAAAGCTCTTGACTCTGTGAGCCGAGTGTAACTTTAAACCCCA 904

QY 958 ATAAACGGATACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACC 1017
Db 905 ACAGCGCATACAGTAGAGGAGCACTGGCTCACCCCTTACCTGGAAACAGTACTATGATC 964

QY 1018 CGACGGATAGACGAGTGGCGGAGGCGCTTACCTTCCGCGATGGAGCTGGATGACCTAC 1077
Db 965 CGACATGAACCAAGTGGCTGGAGGAGCACTTACCTTTGACATGGAGCTGGATGATCTCC 1024

QY 1078 CTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGC 1137
Db 1025 CCAGGAGCGGCTGAAGGAGCTGATCTTCCAGAGACAGCCGCTTCCAGCCAGGGGAC 1084

QY 1138 TGGAGGCCCTTAGCCGACAGACAGACATCTGTGACCCCTGGGCGCTGGAAACCTGGCA 1197
Db 1085 CAGAGGCCCTTAAAGAAACAGACACCCCTGTCTTTTGGACCTGG----- 1131

QY 1198 AAGAGCAAGAGTCACTAGGCGCTCTGTACCCAGAGCCTGCTCTGCTGCGCTGCCCTC 1257
Db 1132 -----TCTGCTCTACTGCTCT 1149

QY 1258 TCCCGCCAGACTGTTAGAAAATGGACACTGTGCGAGCCCGGACCTTGGCAGCCAGGCG 1317
Db 1150 TCTCTGCAAGTGTAGAAAATG--AACTTGTCTAAACCGGACCCCGGACCCAGG-C 1206

QY 1318 GGGGTGGAGATGGGCTGGCCACCTCTCTCTCTTTGTCTGAGGCGCTCCAGCTTCAGCAGG 1377
Db 1207 TGGACCAAGGTGGGCTGGCACTCTCTCACTGTGCTGGGCTCTCTCTGTTCAAG--AG 1264

QY 1378 CCAAGGCCCTTCTCTCCCAACCGGCTTCCCAAGGGGCTCGGAGCTCAGGTGGCCCT 1437
Db 1265 GCTTCTCCCACTCCAGTCCGCTGCCCATCTCCCTTTGACCTGAGTGTGATGGTGTCTCC 1324

QY 1438 AGTTCAATCTCCGCTGCTGCTGCGGCGCTTACCTTCCCGAGCTCCAGTCTCTGGC 1497
Db 1325 AGAGCTGATCTCTGCTGCTG-TGTCTTTATCTATCTCTGCTAGCCCACTCTGTTAGAC 1383

QY 1498 AGTTCTGGAATGGAAGGCTTCTGGCTGCCCACTGCTGAAAGGCGAGAGTGGAGGGTG 1557
Db 1384 GGTCTGGAATGGAAGGCTATGACCGCCTA-----GGACTGTGCTACAGAGGGGTG 1437

QY 1558 GGGGGCGCTGAGTGAAGGATCTAGGCGCATGCTTGGCCCTCATCTCATTTCAAAACCCCA 1617
Db 1438 GAGGGCACTGAGT-----AGGTAAGCTCTGCGCTACTCATCTGTGTAACCCCA 1489

QY 1618 CCTAGTTTCCCTGAAGAACATCTTAGTCTCAAGGGCTGACATCCCTGAGGAGCCAGG 1677
Db 1490 CCAATTTTCCCTGACAGAACATCTTAAATCTCAAGGGCTAGTTTCCCTGAGGAGCCAG 1549

QY 1678 CCGGGCGGAATCCCTCTCCCTGTCAAAAGCTGTCACTTGGGCTGCCCTCGCTGCTCTGTGT 1737
Db 1550 CTAGGCTTAACCTCTCTCTCTC-AAAGTGGCACATGTAAGCCCTTGTGCTTCTGTGT 1608

QY 1738 GT-GGTGAGCAGAAATGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTTGCACCTCCCT 1796
Db 1609 GTGGGTGATGGATGTGGAGGCGGGCGCGTGGAGAGCCCGTGGCGCCCTTCCCTCCCT 1668
```

```
QY 1797 GACCC---GTCTAATATATAAATATAGAGATGTGTCTATGG 1834
Db 1669 GTGCGTGTATCTAATATATAAATATAGAGATGTGTATATGG 1709

RESULT 9
US-08-932-012C-1
; Sequence 1, Application US/08932012C
; Patent No. 6297035
; GENERAL INFORMATION:
; APPLICANT: Boulton et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; FILE REFERENCE: REG 430-Y-1
; CURRENT APPLICATION NUMBER: US/08/932,012C
; CURRENT FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 08/462,874
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 08/178,488
; PRIOR FILING DATE: 1994-01-07
; PRIOR APPLICATION NUMBER: 07/701,544
; PRIOR FILING DATE: 1991-05-16
; PRIOR APPLICATION NUMBER: 07/532,004
; PRIOR FILING DATE: 1990-06-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: RAT
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1095)
; OTHER INFORMATION: ERK1 cDNA
US-08-932-012C-1

Query Match 55.8%; Score 1025; DB 3; Length 1747;
Best Local Similarity 78.1%; Pred. No. 2.1e-202;
Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

QY 58 GNACGAGGGGTCCGCGGGGTCCGCGGGAGGTGGAGATGTTGAAGGGCAGCCGT 117
Db 5 GAACTGCTGGGTCTGTCGCGGTGTCGCGGGAGGTGGAGTGTGAAGGGCAGCCAT 64

QY 118 TCACGCTGGGCGCGGCTACACGCACTGTCAGTTCAGTACATCGCGAGGGCGGTACGGCATGG 177
Db 65 TCGAGTGGGCCCAAGCTTACACGCACTGTCAGTACATCGCGAGGGCGGTACGGCATGG 124

QY 178 TCAGTCTGGCTTATGACCAAGTGGCAAGACTCGGTGGCCATCAAGAGATCAGGCCCT 237
Db 125 TCAGCTCAGCATATGACCAAGTGGCAAGACTCGGTGGCCATCAAGAGATCAGGCCCT 184

QY 238 TCGAACATCAGACCTACTGCGCAGCCGACGCTCGGGAGATCCAGATCTTGGCTGGCTTCC 297
Db 185 TCGAGCTCAAAACCTACTGTCAGCGCACGCTGAGAGAAATCCAGATCTTGGCTGGATTC 244

QY 298 GCATGAGATGTCATCGGCATCCGAGACATCTCGGGCGTCCACCTTGGAGGCCATGA 357
Db 245 GCATGAGATGTCATAGGCATCCGAGACATCTCAGAGCACCCCTTGGAGGCCATGA 304

QY 358 GAGATGTCATATTGTGAGGACCTGATGGAGACTGACTGTACAAAGTTGCTGAAAAGCC 417
Db 305 GAGATGTTTACATTGTTTCAAGACCTCATGGAGACGCACTGTGTAAGCTGCTTAAAGAGCC 364

QY 418 AGCAGTGAAGCAATGACCAATATCTCTACTTCTACAGATCTTGGGGGCTCAAGT 477
Db 365 AGCAGTGAAGCAATGACCAATATCTCTACTTCTCTACAGATCTTGGGGGCTCAAGT 424

QY 478 ACATCCACTCCGCAACGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACA 537
Db 425 ACATACATCTCGGCCAATGTGCTGCAACCGGAGCTGAAAGCCCTCCAATCTGCTTATCAACA 484

QY 538 CCACCTGCGACCTTAAGATTTTGTGATTTTGGGCTCGCCCGGATTCGCGATCTTGAGCATG 597
Db 538 CCACCTGCGACCTTAAGATTTTGTGATTTTGGGCTCGCCCGGATTCGCGATCTTGAGCATG 597
```

Db 485 CCACCTGCGACCTTAAGATCTGTGATTTTGGCCCTGTCGCCGGATGTCGAGCCCTGAGCAG 544
Qy 598 ACCACACCGGCTTCTGACCGAGTATGTGGCTACGGCTGCTGACCGGGCCCCAGAGATCA 657
Db 545 ACCACACTGGCTTCTGACCGAGTATGTGGCCACACGCTGTTACCGAGCCCCAGAGATCA 604
Qy 658 TGTGTAATCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGTGGGCTGCAATCTGG 717
Db 605 TGTCTAACTCAAGGGCTACACCAAAATCCATTCGATCTGCTGTGTGGGCTGCAATCTGG 664
Qy 718 CTGAGATGCTCTTAACCGGGCCATCTTCCCTGCGAAGCACTACCTGCGATCAGCTCAACC 777
Db 665 CTGAGATGCTCTTAACCGGGCTATCTTCCCGCGAAGCACTACCTGCGAAGCTCAACC 724
Qy 778 ACATTTCTGGGCATCTCGGGCTCCCAATCCAGGAGGACCTGAAATTTGTATCATCAACATGA 837
Db 725 ACATTTAGGTATACTGGGTTCCCATCCCAAGGAGCTAAATTTGTATCATCAACATGA 784
Qy 838 AGGCCCGAAACTACTTACATCTCTGCTCCCTCCAAAGCAAGGTGGCTTGGGCCCAAGCTTT 897
Db 785 AGGCCCGAAACTACTTACATCTCTGCTCCCTCTAAACCAAGGTGGCTTGGGCCCAAGCTTT 844
Qy 898 TCCCCAAGTCAAGTCCAAAGCCCTTGACCTGCTGACCGGATGTTAAACCTTTAAACCCCA 957
Db 845 TTCCCAANTCTGATCTCCAAAGCTCTTGACCTGCTGGACCGGATGTTAAACCTTTAAACCCAA 904
Qy 958 ATAAACGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACC 1017
Db 905 ACHAGCGCATCACATGATAGGAGACATGGCTCACCTTACCTGGACAGTACTATGATC 964
Qy 1018 CGACGGATGACCCAGTGGCCGAGAGCCCTTCACTTTCGCGCATGGAGCTGGATGACCTAC 1077
Db 965 CGACAGATGAACAGTGGCTGAGGAGCAATTCACCTTTGACATGGAGCTGGATGATCTCC 1024
Qy 1078 CTAAGGCGGCTCAAGAGCTCATCTTCCAGGAGCAGCAGCTTCCAGCCCGGAGTGC 1137
Db 1025 CCAAGGAGCGGCTGAAGAGCTGATCTTCCAGGAGCAGCCCGCTTCCAGCCAGGGGCAC 1084
Qy 1138 TGGAGGCCCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCCCTGGAAACAGAACTGGCA 1197
Db 1085 CAGAGGCCCCCTAACAGAACAGACACCCCTGTCTCTTTTGGACCTGG----- 1131
Qy 1198 AAGAGGCAAGAGTCACTGAGGGCTCTGTACCCAGGACCTGCTCTCTGCTGCCCCCTC 1257
Db 1132 -----TCTGCTCTACTGCTCTCT 1149
Qy 1258 TCCCGCCAGACTGTTAGAAAATGACACATGTGCCAGCCGGACCTTGGCAGCCAGGCC 1317
Db 1150 TCTCTGAGATTTGTTAGAAAATG--AACTTTGCTCAACCCGGACCCCGGAGCCAGG-C 1206
Qy 1318 GGGGTGGAGCATGGCCCTGGCCACCTCTCTCTTTTGTGTAGGCTTCCAGCTTTCAGGCAGG 1377
Db 1207 TGGACCAAGGGTGGGCTTGGACCCCTCTCACTGTCTGGGGTCTCTCTGTTCAAG--AG 1264
Qy 1378 CCAAGGCTTCTCTCTCCCAACCGCCCTTCCCAACGGGGCTCTGGGAGCTCAGTGGCCCC 1437
Db 1265 GCTTCTCCCATCTCCAGTCCCTGCCCCCATCTCCCTTTGACCTGATGATGAGTGGTCCC 1324
Qy 1438 AGTTCAATCTCCCGCTGTGCTGTGCGCCCTTACCTTCCCAGCGTCCCAAGTCTCTGGC 1497
Db 1325 AGAGCTGATCTCTGCTGTG--TGTCTTTATCTATCTCTCTGCTAGCCCCAGCTCTGCTAGAC 1383
Qy 1498 AGTTCTGGAATGGAAGGGTCTGCTGCTGCCCAACCTGTCTGAAGGCGAGAGTGGAGGGTG 1557
Db 1384 GGTCTGGAATGGAAGGGCTATGACCGCCCTA-----GGACCTGTGCTACAGAGGGGTG 1437
Qy 1558 GGGGGCGCTGAGTAGGGACTCAGGGCCAATGCTTCCCGCCCTCATCTCATTTCAAAACCCCA 1617
Db 1438 GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCTTACTCATCTCTTGTGGAACCCCA 1489
Qy 1618 CCTAGTTTCCCTGAAGGAACTTCTTGTCTCAAGGGCTAGCATCTCTGAGGAGCCAGG 1677
Db 1490 CCCATTTTCCCTGACAGAACTTCTTAATCTCAAGGGCTAGTTTTCCTCTGAGGAGCCAGC 1549

Qy 1678 CCGGGCGGAATCCCTCTCCCTGTCAAAGCTGTCACTTCGGCTGCCCTCGCTGCTCTGTGT 1737
Db 1550 CTAGGCGCTAACCCCTCTCCCTCTC-AAAGCTGCCACATGTAACGCCCTTGTCTTCTGTGT 1608
Qy 1738 GT-GGTGAGCAGAAAGTGGAGCTGGGGGGCTGAGAGCCCGGGCGCCCTGCCACCTCCCT 1796
Db 1609 GTGGGTGATTTGATGTGGAGCGGGGCCCTGTGGAGAGCCCGTGGCCCTCCCACTCCCT 1668
Qy 1797 GACCC---GTCTAATATATAAATATATAGAGATGTGTCTATGG 1834
Db 1669 GTGCTGTATCTAATATATAAATATATAGAGATGTGTATATGG 1709

RESULT 10
US-08-888-818C-1
; Sequence 1, Application US/08888818C
; Patent No. 6303358
; GENERAL INFORMATION:
; APPLICANT: Boulton et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; FILE REFERENCE: REG 430-V-1
; CURRENT APPLICATION NUMBER: US/08/888,818C
; CURRENT FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: 08/478,985
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/178,488
; PRIOR FILING DATE: 1994-01-07
; PRIOR APPLICATION NUMBER: 07/701,544
; PRIOR FILING DATE: 1991-05-16
; PRIOR APPLICATION NUMBER: 07/532,004
; PRIOR FILING DATE: 1990-06-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: RAT
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1095)
; OTHER INFORMATION: ERK1 cDNA
US-08-888-818C-1

Query Match 55.8%; Score 1025; DB 3; Length 1747;
Best Local Similarity 78.1%; Pred. No. 2,1e-202;
Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

Qy 58 GAACCCGAGGGGTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCGAGCCGT 117
Db 5 GAACTGTCTGGGTCGTCTCCCGTGTCTCCCGGGAGGTGGAGGTGGTGAAGGGCGAGCCAT 64
Qy 118 TCACAGTGGGCCCCCGCTACACGAGTTGCAGTACATCGCGAGGGCGCGTACCGCATGG 177
Db 65 TCACAGTGGGCCCCCGCTACACGAGTTGCAGTACATCGCGAGGGCGCGTACCGCATGG 124
Qy 178 TCAGCTCGGCTATGACCACTGCGCAAGATCTGCGTGGCCATCAAGAGATCAGCCCT 237
Db 125 TCAGCTCAGCATATGATCCACGTCGCGCAAGACCAAGTGGCTATCAAGAAGATCAGCCCT 184
Qy 238 TCGAACATCAGACCTACTGCGGCGACGCTCCGGGAGATCCAGATCCTGTCGCTTCC 297
Db 185 TCAGAGCATCAAACTACTGTCTGCGCAAGTCTGCGGAGAAATCCAGATCTTGTCTCGATTC 244
Qy 298 GCCATGAGAAATGTCATCGGATCCGAGACATCTTCGGGGGTCCACCTCGAAGCCATGA 357
Db 245 GCCATGAGAAATGTCATAGGATTCGAGACATCTCTCAGAGCACCCACCTCGAAGCCATGA 304
Qy 358 GAGATGTCTACATTTGTGTCAGGACCTGATGGAGACTGACCTGTGTACAGTTGCTGAAAAGCC 417
Db 305 GAGATGTTTACATTTGTTTCAGGACCTCATGAGAGCGGACCTGTACAGCTGCTAAAGAGCC 364
Qy 418 AGCAGCTGAGCAATGACCAATATCTGCTACTTCTTACCAGATCTCTCGGGGGCTCAAGT 477

Db 365 AGCAGCTGAGCAATGACCAATCTGCTACTTCTCTTACAGATCTCTCCGGGCTCAAGT 424
Qy 478 ACATCACTCCGCGCAACGCTGCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACA 537
Db 425 ACATCACTCGGCAATGTGCTGACCGGACCTGAAGCCCTCCAATCTGCTTATCAACA 484
Qy 538 CCACCTGCGACCTTAAGATTGTGATTTCGGCTGCGCCGGATGCGGATCCGTGAGCATG 597
Db 485 CCACCTGCGACCTTAAGATTGTGATTTCGGCTGCGCCGGATGCGGATCCGTGAGCATG 544
Qy 598 ACCACACCGCTTCCTGACGAGTATGTGGCTACGCGCTGTACCGGGCCCGAGAGATCA 657
Db 545 ACCACACTGGCTTCTGACCGAGTATGTGGCTACGCGCTGTACCGGGCCCGAGAGATCA 604
Qy 658 TGTGTAATCTCAAGGCTATACCAAGTCCATCGACATCTGTGCTGTGGGCTGTGATCTGG 717
Db 605 TGTAACTCCAAGGCTACACCAATCCATTGACATCTGGTCTGTGGGCTGTGATCTGG 664
Qy 718 CTGAGATGCTCTTAACCGGCCATCTCCCTGGCAAGCACTACCTGGATCAGCTCAACC 777
Db 665 CTGAGATGCTCTCAACCGGCCATCTCTCCCGCAAGCACTACCTGGACCAAGCTCAACC 724
Qy 778 ACATTCTGGGCTATCTGGGCTCCCATCCAGGAGCACTGAAATTGTATCATCAACATGA 837
Db 725 ACATTCTAGGTACTGGGTTCCCATCCCAAGAGGACCTAAATTGTATCATTAACATGA 784
Qy 838 AGCCCGAACTACCTACAGTCTCTGCTCCCTCCAGACCAAGTGGCTTGGGCAAGCTTT 897
Db 785 AGCCCGAACTACCTACAGTCTCTGCTCCCTTAAACCAAGTGGCTTGGGCAAGCTTT 844
Qy 898 TCCCAAGTCAAGTCCCAAGCCCTTGAAGTCTGCTGACCGGATGTTAACTTTAACCCCA 957
Db 845 TTTCCCAATCTGACTCCAAAGTCTTTGACTCTGTGACCGGATGTTAACTTTAACCCCA 904
Qy 958 ATAAACCGATCACAGTGGAGGAGCCCTGGCTCACCCCTACCTGGAGCACTACTATGACC 1017
Db 905 ACNAGCGCATCACAGTAGGAGGACACTGGCTACCCCTTACCTGGAGCACTACTATGATC 964
Qy 1018 CGACGATGAGCAGTGGCGGAGGAGCCCTTCACTTTCGCTGAGAGTGGATGAGCTTAC 1077
Db 965 CGACAGATGAACAGTGGCTGAGGAGGACCTTCACTTTGACATGGAGCTGGATGATCTCC 1024
Qy 1078 CTAAAGAGCGGCTGAAGGAGCTCATCTTCCAGAGACAGACGCTTCCAGCCGCGAGTGC 1137
Db 1025 CCAAGAGCGGCTGAAGGAGGCTGATCTTCAAGAGACAGCCGCTTCCAGCCAGGAGGAC 1084
Qy 1138 TGGAGGCCCTTAGCCCGACAGACATCTCTGACCCCTGGGCTTGGAAACAGAACTGGCA 1197
Db 1085 CAGAGGCCCTTAACAGAACAGACACCCCTGTCTTTTGGACCTGG----- 1131
Qy 1198 AAGAGCAAGAGTCACTGAGGCGCTGTGTACCCAGGACCTGCTCTGCTGCTGCCCTTC 1257
Db 1132 -----TCTGCTCTACCTGCTCCT 1149
Qy 1258 TCCCGCCAGACTGTTAGAAAATGGAACATGTGTCACCGCCGCGGACCTTGGCGCCAGGCC 1317
Db 1150 TCTCTGAGATTGTTAGAAAATG- -AATTTGTCTCAACCCGAGCCCGGAGCCCAAGG- 1206
Qy 1318 GGGGTGGAGCATGGGCTGGCCACCTCTCTCTCTTTGCTGAGGCGCTCCAGCTTCAGGCAG 1377
Db 1207 TGGACCAAGGTTGGGCTGGACCCCTCTCACTGTGCTGGGCTCTCTCTGTTCAAG- -AG 1264
Qy 1378 CCAAGGCTTCTCTCCCAACCGCCCTTCCCAAGGGGCTCGGGAGCTCAGGTGCGCCC 1437
Db 1265 GCTTCTCCCACTCCAGTCCCTCTGCCCACTCTCCCTTTGACCTGAGTGAAGGTGGTCCC 1324
Qy 1438 AGTTCAATCTCCGCTGCTGCTGCGCCCTTACCTTCCCAAGCGTCCCAAGTCTCTGAG 1497
Db 1325 AAGAGTGATCTCTGCTGCTG- -TGTCTTATCTATCTCTGCTAGCCCAAGCTCTGTTAGAC 1383
Qy 1498 AGTTCTGGAAATGGAAGGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGTTG 1557

Db 1384 GGTTCTCGAATGGAAGGGCTATGACGCCCTA-----GGACCTGTGTCTACAGAGGGGTG 1437
Qy 1558 GGGGGGCTGAGTAGGGACTCAGGGCCATGCTGCTGCCCTCATCTCATTTCAAAACCCAC 1617
Db 1438 GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCTACTCATCTGTTGGAACCCAC 1489
Qy 1618 CCTAGTTTCCCTGGAAGGAACATTCTTACTCTCAAGGGCTAGCATCCCTTGAGAGCCAGG 1677
Db 1490 CCATTTTCCCTGACAGAACATCTCTAAATCTCAAGGGCTAGTTTCCCTTGAGAGCCAGC 1549
Qy 1678 CCGGCGCAATCCCTCTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 1737
Db 1550 CTAGGCTTAACTCTCTCTCTC-AAAGTGCACATGTAAACGCTTGTGCTTCTGCT 1608
Qy 1738 GT-GGTGACAGAACTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTGCCACCTCCCT 1796
Db 1609 GTGGGTGATTTGATGTGGAGGGGGGGCGCTGGAGAGCCCGTGGCCCTCCCCACCTCCCT 1668
Qy 1797 GACCC---GTCTAAATATATAATATAGAGATGTCTATGG 1834
Db 1669 GTGCTGTATCTAAATATATAATATAGAGATGTATATGG 1709

RESULT 11
US-09-417-197-56
; Sequence 56, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
; FILE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: merKL-F64L-S65T-GFP fusion
; NAME/KEY: CDS
; LOCATION: (1)..(1872)
US-09-417-197-56

Query Match 52.2%; Score 958.6; DB 3; Length 1875;
Best Local Similarity 90.0%; Pred. No. 1.1e-188;
Matches 1027; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 14 GCGGCGGCGGCGGCTCAGGGGGGGGGGGCGGGGAGCCCGTAGAACCGAGGGGGTGG 73
Db 6 GCGGCGGCGGCGGCTCAGGGGGGGGGGGCGGGGAGCCAGGGGAACCTGCTGGGGTCTG 65
Qy 74 CCGGCGGCTCCCGGGGAGTGGAGATGCTGAAGGGGAGCGGCTTCAGAGTGGGCGCCGG 133
Db 66 CCGGCTGCTCCCGGGGAGTGGAGGTTGTTGAAGGGGAGCCATTCGATGTGGGCCACG 125
Qy 134 CTACACGCAAGTTGCAATACATCGGCGAGGGCGCTGACGGCATGGTCAGCTCGGCTATGA 193
Db 126 CTACACGCAAGTGCAGTACATCGGCGAGGGCGCTACGGCATGGTCAGCTCAGCTTATGA 185
Qy 194 CCAGTGGCGAGACTCCGGTGGCCATCAAGAGATCAGCCCTTCGACATCAGACCTA 253
Db 186 CCAGTGGCGAAGCAGAGTGGCCATCAAGAGATCAGCCCTTTGAGCATCAAACTA 245
Qy 254 CTGCCAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGCTCCGCGCATGAGAAATGCTAT 313
Db 246 CTGTACGCGCAGCTGAGGAGATCCAGATCTTGTGCTGCTCCGCGCATGAGAAATGCTAT 305
Qy 314 CCGCATCCGAGACATTTCTGCGGGGCTCCACCTCGAAGCCATGAGAGATGCTTACATTGT 373

306 AGGCATCCGAGACATCTCCAGAGCGCCACCCCTGGAAAGCCATGAGAGATGTTTACATTTGT 365
 374 GCAGGACCTGATGAGAGTCACTGTACAAAGTTGCTGAAAGCCAGCAGCTGAGCAATGA 433
 366 TCAGGACCTCATGAGACAGACCTGTACAAAGCTGCTTAAAGCCAGCAGCTGAGCAATGA 425
 434 CCATATCTGCTACTTCTTACAGATCTCTGCGGGGCTCAAGTATACATCCACTCCGCGAA 493
 426 CCATATCTGCTACTTCTTACAGATCTCTCCGGGCTCCAGTATATATACATCCAGCGAA 485
 494 CTTGCTCCAGGAGATCTAAAGCCCTCAACCTGCTCATCAACACACACCTGCGACCTTAA 553
 486 TGTGCTGCACCGGACCTGAAGCTTCCAACTGCTTATCAACACACACCTGCGACCTTAA 545
 554 GATTGTGATTTTGGGCTGCGGATTCGCGATCTCTGAGCATGACACACACCGGCTTCCT 613
 546 GATCTGTGATTTTGGGCTGCGGATTTGCTGACCGACGACACACACACTGGCTTCT 605
 614 GACGAGTATGTGGCTAGCGCTGGTATACCGGGCCCAAGATCATGCTGAACCTCCAGGG 673
 606 GACGAGTATGTGGCCACAGCTGGTATACCGAGCCCAAGATCATGCTTAAATCCAGGG 665
 674 CTATACCAAGTCCATGACATCTGCTGTGGCTGCAATCTGCTGAGATGCTCTCTAA 733
 666 CTACACCAATCCATGACATCTGCTGTGGCTGCAATCTGCTGAGATGCTCTCTCAA 725
 734 CCGGCCCATCTTCCCTGGCAAGCACTACTGGATGATCAACACACATCTGCGGCACTCT 793
 726 CCGGCCCATCTTCCCGGCAAGCACTACTGGACGCTCAACACACATCTTAGGTATCTT 785
 794 GGGCTCCCATCTCCAGGAGCACTGAATTTGATCATCAACATGAAGCCCGGAACTACCT 853
 786 GGGTTCCCATCTCCAGGAGCACTTAATTTGATCATCAATTAACATGAAGCCCGGAACTACCT 845
 854 ACAGTCTCTGCTCCCAAGACCAAGTGGCTGGCTGGGCAAGCTTTTCCCAAGTCAGACTC 913
 846 GCAGTCTCTGCTCCGAAACCAAGTGGCTGGCTGGGCAAGCTTTTCCCAAGTCAGACTC 905
 914 CAAGACCTTGGCTGCTGGACGGATGTTAACTTTAACTCCCAATGAAGCAAGTCAAGT 973
 906 CAAGCTCTTGGCTGCTGGACGGATGTTAACTTTCACTCCCAAGCAAGGCGATCAGT 965
 974 GGAGGAAGCGTGGCTCACTTACCTGAGCAGTACTATGACCCGAGGATGAGCCAGT 1033
 966 AGAGGAAGCGTGGCTCACTTACCTGGAACAGTACTAGATCCGACAGATGAGCCAGT 1025
 1034 GCGCGAGGAGCTTCACTTCCGATGAGCTGAGTGAATGACCTTAAAGGAGCGGCTGAA 1093
 1026 GCGCGAGGAGCAATCACTTCCGATGAGCTGAGTGAATGACCTTCCCAAGGAGCGGCTGAA 1085
 1094 GGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCGGAGTGTGAGGCGCCCTAGCC 1153
 1086 GGAGTGAATCTTCCAGGAGACAGCCGCTTCCAGCGGCGCCAGAGGCGCCCGGGCG 1145
 1154 C 1154
 1146 C 1146

RESULT 12
 US-09-949-016-4048
 ; Sequence 4048, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4048
 ; LENGTH: 1701
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-4048

Query Match 35.5%; Score 651.4; DB 3; Length 1701;
 Best Local Similarity 74.4%; Pred. No. 2.8e-125;
 Matches 819; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

33 GGGGGCGGGGGGGGGGGGGGGGGTGTAGAACCGAGGGGGTCCGGCCCGGGGGTCCGGGGGAG 92
 170 GGGGAGGG 229
 93 GTGGAGATGTTGAAGGGGGGAGCGGTTGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 152
 230 CCGGAGATGTTCCGGGGGGGAGGGTGTTCGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGG 289
 153 ATCGGGGAGGG 212
 290 ATCGGGGAGGG 349
 213 GTGGCCATCAAGAAGATCAAGCCCTTTCGAAATCATCAGACCTTACTGCGAGGGGGGGGGGG 272
 350 GTAGCTATCAAGAAGATCAAGCCCTTTCGAGACCAAGGGGGGGGGGGGGGGGGGGGGGG 409
 273 GAGATCCAGATCTGCTGG 332
 410 GAGATAAAGATCTTACTGCGCTTCAGACATGAGAAATCATTTGGAATCAATGACATTTATT 469
 333 CGGGCGTCCACCTTGGAGCCATGAGAGATGTTCTACATTTGTGAGGAGCCCTGATGGAGACT 392
 470 CGAGCACCAACCAATCGAGCAAAATGAAGATGTATATATAGTACAGGACCTCATGGAAACA 529
 393 GACCTGTACAAAGTGTCTGAAAAGCCAGCAGCTGAGCAATGACCAATATCTGCTACTTCTTC 452
 530 GATCTTTACAAAGCTTCTTGAAGACACAACTTCAAGCAATGACCAATATCTGCTATTTTCTC 589
 453 TACAGATCTGCGGGGGCTTCAAGTATCATCTCCGCGCAAGCTGCTCCAGGAGATCTA 512
 590 TACAGATCTCAGAGGGGTTAAAGATATATCCATTCAGCTAACGTTCTGCAACCGGAGCTC 649
 513 AAGCCCTCCAACTGCTCATCAACACCACTGCGAGCTTAAGATTTGTGATTTTGGGCTG 572
 650 AAGCCCTTCCAACTGCTGCTCAACACCACTGCTGATCTCAAGATCTGAGCTTTGGGCTG 709
 573 GCGCGGATTCGCGATCTGAGCATGACCAACCGGGCTTCTGAGGAGTATGTGGCTTACG 632
 710 GCGCGTGTTCAGATCCAGACCATGATCAACAGGGGTTCTGACAGATATGTGGCCACA 769
 633 CGCTGTACCGGG 692
 770 CGTTGGTACAGGGGCTCCAGAAATTAATGTTGAAATTCAGAGGGGCTACACCAAGTCCATTGAT 829
 693 ATCTGCTGTGGGGTGCATTTCTGGCTGAGATGCTCTTAACCGGGGGGGGGGGGGGGGGGG 752
 830 ATTTGGTCTGTAGGCTGCAATTTCTGGCAGAAATGCTTTTAAAGGGGGGGGGGGGGGGGG 889
 753 AAGCAGCTTACCTGGATCAGCTCAACCAATCTGCGGATCTGCGGATCTTGGGCTCCCATCCAGGAG 812
 890 AAGCATTTCTTACAGCAGCTGAACCAATTTTGGTATTTCTTGGATCCCATCACAAGAA 949
 813 GACCTGAATTTGATCATCAACATGAAGCCCGGAAATACCTACAGTCTCTGCGGCTCAAG 872
 950 GACCTGAATTTGATTAATAAATTTAAAGAGCTAGGAACTATTTGCTTTCTTCCACACAAA 1009
 873 ACCAAGTGGCTTGGGCGAAGCTTTTCCCAAGTCTCAGACTCCAAAGCCCTTGACCTGCTG 932

[illegible]

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:49:28 ; Search time 1578.19 Seconds
(without alignments)
9625.515 Million cell updates/sec

Title: US-10-623-108-5
Perfect score: 1837
Sequence: 1 gagagtgagatggcgcg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:
1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgm2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgm2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgm2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1837	100.0	1837	8	US-10-623-108-5
2	1707	92.9	1777	8	US-10-623-108-7
3	1679.8	91.4	1869	5	US-10-098-841-114
4	1595	86.8	1726	8	US-10-623-108-3
5	1451	79.0	1654	8	US-10-623-108-1
6	1285.8	70.0	1400	10	US-11-060-756-1050
7	1285.8	70.0	1400	10	US-11-060-756-5322
8	1138.4	62.0	1896	5	US-10-072-036-38
9	1076	58.6	1214	3	US-09-925-237-323
10	1043.2	56.8	1147	7	US-10-114-270-109
11	958.6	52.2	1875	5	US-10-072-036-56
12	646.6	35.2	1611	5	US-10-171-311-126
13	646.6	35.2	1611	6	US-10-301-822-114
14	642	34.9	2791	6	US-10-305-720-1263
15	642	34.9	2791	9	US-10-756-149-513
16	632.4	34.4	1074	9	US-10-451-291B-11
17	632.4	34.4	1815	5	US-10-072-036-58
18	632.4	34.4	1815	5	US-10-072-036-40
19	603.4	32.8	2111	4	US-09-925-065A-75282
20	603.4	32.8	13808	3	US-09-764-877-3351
21	603.4	32.8	13808	6	US-10-242-515-3351
22	525.8	28.6	546	3	US-09-738-973-457
23	525.8	28.6	546	3	US-09-854-133-457

C	24	525.8	28.6	546	5	US-10-144-649A-457	Sequence 457, App
	25	494	26.9	574	3	US-09-738-973-456	Sequence 456, App
	26	494	26.9	574	3	US-09-854-133-456	Sequence 456, App
	27	494	26.9	574	5	US-10-144-649A-456	Sequence 456, App
	28	433.8	23.6	682	6	US-10-264-049-1992	Sequence 1992, Ap
	29	416.6	22.7	641	7	US-10-404-460-74	Sequence 74, Appl
	30	398.2	21.7	544	3	US-09-918-995-2207	Sequence 2207, Ap
	31	348.8	19.0	463	3	US-09-318-995-22034	Sequence 22034, A
	32	338.8	18.4	1396	8	US-10-768-886-1	Sequence 1, Appl1
	33	336	18.3	400	3	US-09-960-352-572	Sequence 572, App
	34	325.4	17.7	1610	7	US-10-767-701-14213	Sequence 14213, A
	35	320.4	17.4	1945	7	US-10-425-114-18739	Sequence 18739, A
	36	320.4	17.4	2828	6	US-10-305-720-1458	Sequence 1458, Ap
	37	320.4	17.4	2865	7	US-10-425-114-26261	Sequence 26261, A
	38	320.4	17.4	2980	8	US-10-737-450-105	Sequence 105, App
	39	320.4	17.4	2980	8	US-10-723-860-2897	Sequence 2897, Ap
	40	320.4	17.4	3195	7	US-10-425-114-26224	Sequence 26224, A
	41	320.4	17.4	3309	8	US-10-723-860-6956	Sequence 6956, Ap
	42	316.2	17.2	2187	8	US-10-278-698-185	Sequence 185, App
	43	316.2	17.2	2187	8	US-10-278-698-701	Sequence 701, App
	44	312.4	17.0	1502	5	US-10-198-343-1	Sequence 1, Appl1
	45	312.4	17.0	1502	5	US-10-197-315-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

US-10-623-108-5
; Sequence 5, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAL, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCES: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-108-5

Query Match 100.0%; Score 1837; DB 8; Length 1837;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAGAGTGGAGATGGCGCGCGCGCGCTCAGGGGGCGGGGGCGGGGAGCCCGTAGAA	60
Db	1	GAGAGTGGAGATGGCGCGCGCGCGCTCAGGGGGCGGGGGCGGGGAGCCCGTAGAA	60
Qy	61	CCGAGGGGGTCCGCCCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCAGCCGTCG	120
Db	61	CCGAGGGGGTCCGCCCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCAGCCGTCG	120
Qy	121	ACGTGGGCGCGCTACACGAGTTGACATCGCGGAGGGCGCGCTACGGCATGGTCA	180
Db	121	ACGTGGGCGCGCTACACGAGTTGACATCGCGGAGGGCGCGCTACGGCATGGTCA	180
Qy	181	GCTCGGCTTATGACACCGTGGCGAAGACTCGCGTGGGCGCATCAAGAATACGCCCTTCG	240
Db	181	GCTCGGCTTATGACACCGTGGCGAAGACTCGCGTGGGCGCATCAAGAATACGCCCTTCG	240
Qy	241	AACATCAGACTTACTGCGGCGACCGCTCCGGGAGATCCAGATCCTCGTGGCTTCGCC	300
Db	241	AACATCAGACTTACTGCGGCGACCGCTCCGGGAGATCCAGATCCTCGTGGCTTCGCC	300
Qy	301	ATGAGATGTTCATCGGCATCCGAGACATTCTCGGGGGCGTCCACCTTGAAGCCATGAGAG	360
Db	301	ATGAGATGTTCATCGGCATCCGAGACATTCTCGGGGGCGTCCACCTTGAAGCCATGAGAG	360

Qy 361 ATGTCTACATTGTGACGAGCCTGATGAGAGCTGACCTGTACAAGTTGCTGAAAAGCCAGC 420
Db 361 ATGTCTACATTGTGACGAGCCTGATGAGAGCTGACCTGTACAAGTTGCTGAAAAGCCAGC 420
Qy 421 AGCTGACCAATGACCATATCTGCTACTTCTCTACAGATCTCGCGGGCTCAAGTACA 480
Db 421 AGCTGACCAATGACCATATCTGCTACTTCTCTACAGATCTCGCGGGCTCAAGTACA 480
Qy 481 TCCACTCCGCAACGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
Db 481 TCCACTCCGCAACGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
Qy 541 CTTGCGACCTTAAGATTGTGATTTGGCCCTGGCCCGGATTCGCGATCCTGAGCATGACC 600
Db 541 CTTGCGACCTTAAGATTGTGATTTGGCCCTGGCCCGGATTCGCGATCCTGAGCATGACC 600
Qy 601 ACACGGCTTCTGACCGAGTATGTGCTACGGCTGGTACCGGCCCGCCAGAGATCATGC 660
Db 601 ACACGGCTTCTGACCGAGTATGTGCTACGGCTGGTACCGGCCCGCCAGAGATCATGC 660
Qy 661 TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCATTCGGCTG 720
Db 661 TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCATTCGGCTG 720
Qy 721 AGATGCTCTCTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 780
Db 721 AGATGCTCTCTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 780
Qy 781 TTCTGGGATCTCGGGCTCCCATCCAGAGAGGACCTGAATGTATCATCAAGATGAAG 840
Db 781 TTCTGGGATCTCGGGCTCCCATCCAGAGAGGACCTGAATGTATCATCAAGATGAAG 840
Qy 841 CCCGAACCTACTACAGTCTCTGCCCCCTCCAAAGTGGCTTGGGCCAAGCTTTTCC 900
Db 841 CCCGAACCTACTACAGTCTCTGCCCCCTCCAAAGTGGCTTGGGCCAAGCTTTTCC 900
Qy 901 CCAAGTCAGACTCCAAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 960
Db 901 CCAAGTCAGACTCCAAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 960
Qy 961 AACGGATCAGCTGAGGAGAACGGCTGCTACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Db 961 AACGGATCA----- 969
Qy 1021 CGGATGAGCCAGTGGCCGAGGAGCCCTTACCTTTCGCCATGAGCTGGATGACCTACTA 1080
Db 970 -----CAGTGGCCGAGGAGCCCTTACCTTTCGCCATGAGCTGGATGACCTACTA 1020
Qy 1081 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1140
Db 1021 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1080
Qy 1141 AGGCCCCCTAGCCAGACAGACATCTCTGCAACCTTGGGCCCTGGAAAGAACTGGCAAAG 1200
Db 1081 AGGCCCCCTAGCCAGACAGACATCTCTGCAACCTTGGGCCCTGGAAAGAACTGGCAAAG 1140
Qy 1201 AGGCAAGAGTCTAGTGGGCTCTGTACCCAGGACCTGCTCTGCTGCTGCCCTCTCC 1260
Db 1141 AGGCAAGAGTCTAGTGGGCTCTGTACCCAGGACCTGCTCTGCTGCTGCCCTCTCC 1200
Qy 1261 GGCAGAGCTTTAGAAAATGACACTGTGCCCCAGCCCGGACCTTGGCAGCCCGCCGGG 1320
Db 1201 GGCAGAGCTTTAGAAAATGACACTGTGCCCCAGCCCGGACCTTGGCAGCCCGCCGGG 1260
Qy 1321 GTGAGCATGGGCTGGCCACTCTCTCTTTGTGAGGCTTCAGCTTTCAGGAGGCCA 1380
Db 1261 GTGAGCATGGGCTGGCCACTCTCTCTTTGTGAGGCTTCAGCTTTCAGGAGGCCA 1320
Qy 1381 AGGCTTTCTCTCCCAACCCGCTTCCCAAGGGCTCTGGGAGCTCAGGTGGCCCAAGT 1440
Db 1321 AGGCTTTCTCTCCCAACCCGCTTCCCAAGGGCTCTGGGAGCTCAGGTGGCCCAAGT 1380

Qy 1441 TCAATCTCCGCTGCTGCTGCTGCGCCCTTACCTTCCAGAGCTCCAGTCTCTGGCAGT 1500
Db 1381 TCAATCTCCGCTGCTGCTGCTGCGCCCTTACCTTCCAGAGCTCCAGTCTCTGGCAGT 1440
Qy 1501 TCTGGAATGGAAGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1560
Db 1441 TCTGGAATGGAAGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1500
Qy 1561 GGCGCTGAGTGAAGGATCTCAGGGCCATGCTGCGCCCTCATCTCATTTCAAAACCCACCT 1620
Db 1501 GGCGCTGAGTGAAGGATCTCAGGGCCATGCTGCGCCCTCATCTCATTTCAAAACCCACCT 1560
Qy 1621 AGTTTCCCTGAAGGACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCG 1680
Db 1561 AGTTTCCCTGAAGGACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCG 1620
Qy 1681 GGCGGAATCCCTCCCTGTCANAGCTGTCACTTTCGGGTGCCCTCGCTGCTTCTGTGTG 1740
Db 1621 GGCGGAATCCCTCCCTGTCANAGCTGTCACTTTCGGGTGCCCTCGCTGCTTCTGTGTG 1680
Qy 1741 GTGAGCAGAAAGTGGAGCTGGGGGGCTGGAGAGCCCGGCCCTGCGCACCCTCGCTGACC 1800
Db 1681 GTGAGCAGAAAGTGGAGCTGGGGGGCTGGAGAGCCCGGCCCTGCGCACCCTCGCTGACC 1740
Qy 1801 CGTCTAATATATATATATATAGAGATGTCTATGCGTG 1837
Db 1741 CGTCTAATATATATATATATAGAGATGTCTATGCGTG 1777

RESULT 3

US-10-098-841-114
; Sequence 114, Application US/10098841
; Publication No. US2002019769A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US2002019769A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 114
; LENGTH: 1869
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)..(1212)
US-10-098-841-114

Query Match 91.4%; Score 1679.8; DB 5; Length 1869;

Query Match		86.8%; Score 1595; DB 8; Length 1726;
Best Local Similarity		94.0%; Pred. No. 0;
Matches 1726; Conservative		0; Mismatches 0; Indels 111; Gaps 2;
Qy	1	GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGCGGGAGCCCGTAGAA 60
Db	1	GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGCGGGAGCCCGTAGAA 60
Qy	61	CCGAGGGGGTCCGGCGCGCGCGCTCCGGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTG 120
Db	61	CCGAGGGGGTCCGGCGCGCGCGCTCCGGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTG 120
Qy	121	ACGTGGCCCCGCGCTACACGAGTTGTCAGTACATCGCGGAGGGGCGGTACGGCATGTGTC 180
Db	121	ACGTGGCCCCGCGCTACACGAGTTGTCAGTACATCGCGGAGGGGCGGTACGGCATGTGTC 180
Qy	181	GCTCGGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAGAGATCAGCCCTTCG 240
Db	181	GCTCGGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAGAGATCAGCCCTTCG 240
Qy	241	AACATCAGACCTACTGCGCAGCGACGCTCCGGGAGATCCAGATCCTGCTCGCTTCGCGC 300
Db	241	AACATCAGACCTACTGCGCAGCGACGCTCCGGGAGATCCAGATCCTGCTCGCTTCGCGC 300
Qy	301	ATGAGAAATGTCATCGGCATCCGAGACATTCCTGCGGGGCTCCACCTGGAAGCCATGAGAG 360
Db	301	ATGAGAAATGTCATCGGCATCCGAGACATTCCTGCGGGGCTCCACCTGGAAGCCATGAGAG 360
Qy	361	ATGCTACATGTCAGGACCTGATGAGAGTGAACCTGTACAAAGTTGCTGAAAAGCCAGC 420
Db	361	ATGCTACATGTCAGGACCTGATGAGAGTGAACCTGTACAAAGTTGCTGAAAAGCCAGC 420
Qy	421	AGCTGACCAATGACCATATCTGCTACTTCTCTACAGATCCTCGGGGCTCAAGTACA 480
Db	421	AGCTGACCAATGACCATATCTGCTACTTCTCTACAGATCCTCGGGGCTCAAGTACA 480
Qy	481	TCACCTCGGCAACGCTGCTCCAGAGATCTAAAGCCCTCCACCTGCTCATCAACACCA 540
Db	481	TCACCTCGGCAACGCTGCTCCAGAGATCTAAAGCCCTCCACCTGCTCATCAACACCA 540
Qy	541	CCTCGGACCTTAAGATTGATTTGCTGCTGCTGCGCGCGGATGCGGATCCTGAGCATGACC 600
Db	541	CCTCGGACCTTAAGATTGATTTGCTGCTGCTGCGCGCGGATGCGGATCCTGAGCATGACC 600
Qy	601	ACACCGCTTCTGACGAGATATGCTGCTACGCGCTGCTACCGGGCCCGAGAGATCATGC 660
Db	601	ACACCGCTTCTGACGAGATATGCTGCTACGCGCTGCTACCGGGCCCGAGAGATCATGC 660
Qy	661	TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGCTGGGCTGCAATCTGGCTG 720
Db	661	TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGCTGGGCTGCAATCTGGCTG 720
Qy	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAGGACCTACCTGGATCAGCTCAACACA 780
Db	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAGGACCTACCTGGATCAGCTCAACACA 780
Qy	781	TTCTGGGATCCTCGGCTCCCATCCAGAGGACCTGAATTGTATCATCAACATGAAG 840
Db	781	TTCTGGGATCCTCGGCTCCCATCCAGAGGACCTGAATTGTATCATCAACATGAAG 840
Qy	841	CCCGAAACTACTACAGTCTCTGCTCCCAAGACCAAGGTGGCTTGGGCCAAGCTTTTC 900
Db	841	CCCGAAACTACTACAGTCTCTGCTCCCAAGACCAAGGTGGCTTGGGCCAAGCTTTTC 900
Qy	901	CCAAGTGAGACTCAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 960
Db	901	CCAAGTGAGACTCAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 960
Qy	961	AACGGATCACAGTGGAGGAGCGCTGCTACCCCTACCTGAGCAGTACTATGACCGA 1020
Db	961	AACGGATCA----- 969
Qy	1021	CGGATGAGCCAGTGGCGGAGGAGCCCTTACCTTTCGCCATGGAGCTGGATGACCTACTA 1080

RESULT 5
US-10-623-108-1
; Sequence 1, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAL, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-108-1

Db	970	-----CAGTGGCGGAGGAGCCCTTACCTTCCCATGGAGCTGGATGACCTACTCTA 1020
Qy	1081	AGGAGCGGCTGAAGAGTCTATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGG 1140
Db	1021	AGGAGCGGCTGAAGAGTCTATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGG 1080
Qy	1141	AGGCCCTTAGCCACAGACAGATCTCTGACCCCTGGGGCTGGAAACAGACTGGCAGAG 1200
Db	1081	AGGCCCTTAGCCACAGACAGATCTCTGACCCCTGGGGCTGGAAACAGACTGGCAGAG 1122
Qy	1201	AGSCAAGAGGTCACTGAGGGCTCTCTGACCCAGAGGACCTGCTCTGCTGCTGCTGCTCTCC 1260
Db	1123	-----GACCTGCTTCTGCTGCTGCTGCTCTCC 1149
Qy	1261	CGCCAGATGTTAGAAAATGACACTGTGTCACAGCCCGAGACCTTGGAGAGCCAGGCGGG 1320
Db	1150	CGCCAGATGTTAGAAAATGACACTGTGTCACAGCCCGAGACCTTGGAGAGCCAGGCGGG 1209
Qy	1321	GTGGAGATGGGGCTGGCCACCTCTCTCTTCTGCTGAGGCTCCAGCTTCAGGAGGCCA 1380
Db	1210	GTGGAGATGGGGCTGGCCACCTCTCTCTTCTGCTGAGGCTCCAGCTTCAGGAGGCCA 1269
Qy	1381	AGSCCTTCTCTCCACCCAGCCGCTCCACAGGGGCTCGGGAGCTCAGGTGCCCCAGT 1440
Db	1270	AGSCCTTCTCTCTCCACCCGCTCCACAGGGGCTCGGGAGCTCAGGTGCCCCAGT 1329
Qy	1441	TCATCTCCCGCTGCTGCTGCTGCTGCTTACCTTCCACAGCTCCAGCTCTGAGCAGT 1500
Db	1330	TCATCTCCCGCTGCTGCTGCTGCTGCTTACCTTCCACAGCTCCAGCTCTGAGCAGT 1389
Qy	1501	TCGGAATGGAAGGGTTCTGGCTGCCCCAACCTGCTGAGGGGAGAGGTGGGGGG 1560
Db	1390	TCGGAATGGAAGGGTTCTGGCTGCCCCAACCTGCTGAGGGGAGAGGTGGGGGG 1449
Qy	1561	GGCGCTGAGTGAAGGACTCAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db	1450	GGCGCTGAGTGAAGGACTCAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509
Qy	1621	AGTTTCTCCCTGAAGGAACTTCTTAGTCTCAAGGGCTAGCATCCCTGAGAGGAGGCGG 1680
Db	1510	AGTTTCTCCCTGAAGGAACTTCTTAGTCTCAAGGGCTAGCATCCCTGAGAGGAGGCGG 1569
Qy	1681	GGCGGAATCCCTTCTGCTCAAGAGTGTCACTTGGGTGCTGCTGCTGCTGCTGCTGCTG 1740
Db	1570	GGCGGAATCCCTTCTGCTCAAGAGTGTCACTTGGGTGCTGCTGCTGCTGCTGCTGCTG 1629
Qy	1741	GTGAGCAGAGTGGAGCTGGGGGCTGAGAGCGCGGGGCTGAGAGCGCGGCTGCGCTGAC 1800
Db	1630	GTGAGCAGAGTGGAGCTGGGGGCTGAGAGCGCGGGGCTGAGAGCGCGGCTGCGCTGAC 1689
Qy	1801	CGTCTAATATATATAATATAGAGATGTCTATGGCTG 1837
Db	1690	CGTCTAATATATATAATATAGAGATGTCTATGGCTG 1726

Query Match		79.0%; Score 1451; DB 8; Length 1654;
Best Local Similarity		90.0%; Pred. No. 0;
Matches 1654; Conservative		0; Mismatches 0; Indels 183; Gaps 2;
QY	1	GAGAGTGGAGATGGCGGGCGCGGCTCAGGGGGCGGGGCGGGAGCCCCGTAGAA 60
DB	1	GAGAGTGGAGATGGCGGGCGCGGCTCAGGGGGCGGGGCGGGAGCCCCGTAGAA 60
QY	61	CGAGGGGGTCCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCGAGCCGTTCCG 120
DB	61	CGAGGGGGTCCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCGAGCCGTTCCG 120
QY	121	ACGTGGGCGCGCTACACGAGTTCAGTACATCGGCGAGGCGCGTACGGCATGCTCA 180
DB	121	ACGTGGGCGCGCTACACGAGTTCAGTACATCGGCGAGGCGCGTACGGCATGCTCA 180
QY	181	GCTCGGCGCTATGACACGTCGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCCCTTCG 240
DB	181	GCTCGGCGCTATGACACGTCGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCCCTTCG 240
QY	241	AACATCAGACCTACTGCGACGCGCACGCTCGGGAGATCCAGATCTCTGCTGCGCTTCGCGC 300
DB	241	AACATCAGACCTACTGCGACGCGCACGCTCGGGAGATCCAGATCTCTGCTGCGCTTCGCGC 300
QY	301	ATGAGAATGTTCATCGGCATCCGAGACATTCGCGGGCGCTCCACCTGGGAAGCCATGAGAG 360
DB	301	ATGAGAATGTTCATCGGCATCCGAGACATTCGCGGGCGCTCCACCTGGGAAGCCATGAGAG 360
QY	361	ATGTCTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAGCCAGC 420
DB	361	ATGTCTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAGCCAGC 420
QY	421	AGCTGAGCAATGACATATCTGCTACTTCTCTACAGATCTCTGCGGGCGCTCAAGTACA 480
DB	421	AGCTGAGCAATGACATATCTGCTACTTCTCTACAGATCTCTGCGGGCGCTCAAGTACA 480
QY	481	TCCACTCCGCAACGTCCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
DB	481	TCCACTCCGCAACGTCCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
QY	541	CCTGCGACCTTAAGATTGTTGATTTTCGGCGTGGCCCGGATTCGCGATCTGAGCATGACC 600
DB	541	CCTGCGACCTTAAGATTGTTGATTTTCGGCGTGGCCCGGATTCGCGATCTGAGCATGACC 600
QY	601	ACACCGGCTTCTGACGAGATATGGCTACGCGCTGGTACCGGGCCCGAGATCATGC 660
DB	601	ACACCGGCTTCTGACGAGATATGGCTACGCGCTGGTACCGGGCCCGAGATCATGC 660
QY	661	TGAATCCCAAGGCTATACCAAGTCCATCGACATCTGCTGCGGCTGCATTTCTGGCTG 720
DB	661	TGAATCCCAAGGCTATACCAAGTCCATCGACATCTGCTGCGGCTGCATTTCTGGCTG 720
QY	721	AGATGCTCTCTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 780
DB	721	AGATGCTCTCTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 780
QY	781	TTCTGGGCTATCCGCGTCCCGATCCGAGGAGACCTGAAATGTATCATCAATCAAGG 840
DB	781	TTCTG----- 785
QY	841	CCGAAACTACCTACAGTCTCTGCCCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTTTCC 900
DB	786	----- 785
QY	901	CCAAGTCAGATCCAAAGCCCTTACCTGTCGACCGGATGTTAACTTTAACCCCAATA 960
DB	786	-----GCCCTTGACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 828
QY	961	AACGGATCAGTGGAGGAAGCCCTGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
DB	829	AACGGATCAGTGGAGGAAGCCCTGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 888

QY	1021	CGGATGAGCCAGTGGCCGAGGAGCCCTTACCTTCGCCATGGAGTGGATGACTTACCTTA 1080
DB	889	CGGATGAGCCAGTGGCCGAGGAGCCCTTACCTTCGCCATGGAGTGGATGACTTACCTTA 948
QY	1081	AGGAGGGCTGAAGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1140
DB	949	AGGAGGGCTGAAGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1008
QY	1141	AGGCCCCCTAGCCACAGACAGATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAG 1200
DB	1009	AGGCCCCCTAGCCACAGACAGATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAG 1050
QY	1201	AGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGACCTTGCCTCTGCTGCTGCCCTCTCC 1260
DB	1051	-----GGACCTTGCCTCTGCTGCCCTCTCC 1077
QY	1261	CGCCAGACTGTTAGAAAATGGACATGTGTGCCAGCCCGGACCTTGGCAGCCCGAGCCGG 1320
DB	1078	CGCCAGACTGTTAGAAAATGGACATGTGTGCCAGCCCGGACCTTGGCAGCCCGAGCCGG 1137
QY	1321	GTGGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTCAAGCAGGCCA 1380
DB	1138	GTGGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTCAAGCAGGCCA 1197
QY	1381	AGGCTTCTCTTCCACACCCGCTTCCCAAGGGCTTGGGAGCTCAAGTGGCCCCAGT 1440
DB	1198	AGGCTTCTCTTCCACACCCGCTTCCCAAGGGCTTGGGAGCTCAAGTGGCCCCAGT 1257
QY	1441	TCATCTCCGCTGTGTGTGCTGCGGCTTACCTTCCAGGCTCCAGTCTCTGGCAGT 1500
DB	1258	TCATCTCCGCTGTGTGTGCTGCGGCTTACCTTCCAGGCTCCAGTCTCTGGCAGT 1317
QY	1501	TCGTGAATGGAAGGTTCCTGGCTGCCCAACCTGCTGAAGGCGAGAGTGGAGGGTGGG 1560
DB	1318	TCGTGAATGGAAGGTTCCTGGCTGCCCAACCTGCTGAAGGCGAGAGTGGAGGGTGGG 1377
QY	1561	GGCGCTGATGAGGACTCAGGGCCATGCTGCCCTCCCTCATCTCAATCAAAACCCCACT 1620
DB	1378	GGCGCTGATGAGGACTCAGGGCCATGCTGCCCTCCCTCATCTCAATCAAAACCCCACT 1437
QY	1621	AGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGCCG 1680
DB	1438	AGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGCCG 1497
QY	1681	GGCCGAATCCCTCTCTGCTCAAGGTGTCACTTCCGCTGCCCTGCTGCTCTGTGTGTG 1740
DB	1498	GGCCGAATCCCTCTCTGCTCAAGGTGTCACTTCCGCTGCCCTGCTGCTCTGTGTGTG 1557
QY	1741	GTGAGCAGAAAGTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTGCGCACCTCCCTGACC 1800
DB	1558	GTGAGCAGAAAGTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTGCGCACCTCCCTGACC 1617
QY	1801	CGTCTAATATATAAATATAGAGATGTGTATGGCTG 1837
DB	1618	CGTCTAATATATAAATATAGAGATGTGTATGGCTG 1654

RESULT 6

US-11-060-756-1050
; Sequence 1050, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1050
; LENGTH: 1400

Db 121 TAAAGATTGTGATTTGGCCCTGCGCGGATTCGGATCCTGAGCATGACCAACACCGGCTT 180
QY 611 CTGACGAGATATGTGGCTACGCGCTGTGATACGGGGCCGAGAGATCATGCTGAACCTCAA 670
Db 181 CTTGSGGAGYMYTGRMMWCMGSMGSGTACCGGGGCCGAGAGATCATGCTGAACCTCAA 240
QY 671 GGGCTATACCAAGTCCATGACATCTGGTCTGTGGGCTGATTTCTGGCTGAGATGCTCTC 730
Db 241 GGGCTATACCAAGTCCATGACATCTGGTCTGTGGGCTGATTTCTGGCTGAGATGCTCTC 300
QY 731 TAAACGGCCCATCTTCCCTGGCAAGCATACCTGGATCAGCTCAACCAATCTTGGGCAT 790
Db 301 TAAACGGCCCATCTTCCCTGGCAAGCATACCTGGATCAGCTCAACCAATCTTGGGCAT 360
QY 791 CTGGGCTCCCATCCAGAGAGACCTGAATGTATCATCAACATGAAGGCCCGCAACTA 850
Db 361 CTGGGCTCCCATCCAGAGAGACCTGAATGTATCATCAACATGAAGGCCCGCAACTA 420
QY 851 CTTACAGTCTCTGCCCTCCAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGA 910
Db 421 CTTACAGTCTCTGCCCTCCAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGA 480
QY 911 CTCCAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAACCCCAATAAAACGGATCAC 970
Db 481 CTCCAAGCCCTTGAACCTGCTGGACCGGATGTTAACTTTAACCCCAATAAAACGGATCAC 540
QY 971 AGTGGAGGAAGCGCTGCTCAACCTTACCTGGAGCAGTACTATGACCCGACGGATGAGCC 1030
Db 541 AGTGGAGGAAGCGCTGCTCAACCTTACCTGGAGCAGTACTATGACCCGACGGATGAGCC 600
QY 1031 AGTGGCGGAGGAGCCCTTCACTTCCCATGGAGCTGGATGACTTAAAGGAGCGGCT 1090
Db 601 AGTGGCGGAGGAGCCCTTCACTTCCCATGGAGCTGGATGACTTAAAGGAGCGGCT 660
QY 1091 GAAGGAGCTCATCTTCCAGAGACAGCACGCTTCCAGCCCGGAGTGTCTGGAGGCCCTTA 1150
Db 661 GAAGGAGCTCATCTTCCAGAGACAGCACGCTTCCAGCCCGGAGTGTCTGGAGGCCCTTA 720
QY 1151 GCCCAGACAGATCTCTGACCCCTGGGGCTGGAAACAGAACATGGCCAAAGAGGAGG 1210
Db 721 GCCCAGACAGATCTCTGACCCCTGGGGCT----- 752
QY 1211 TCACTGAGGSCCTCTGTACCCAGAGACCTGCCCTCTGCCCTGCCCTCTCCCGCAGACTG 1270
Db 753 -----GAACTGCCCTCTGCTGCCCCCTCTCCCGCAGACTG 789
QY 1271 TTAGAAAATGGACACTGTGCCAGCCGGACCTTGGCAGCCAGGCGGGGTGGAGCATG 1330
Db 790 TTAGAAAATGGACACTGTGCCAGCCGGACCTTGGCAGCCAGGCGGGGTGGAGCATG 849
QY 1331 GGCTGGCCACCTCTCTCTTTGCTGAGGGCTCCAGCTTCAGGAGCCGAGCCGAGCTTCTC 1390
Db 850 GGCTGGCCACCTCTCTCTTTGCTGAGGCTCCAGCTTCAGGAGCCGAGCCGAGCTTCTC 909
QY 1391 CTCGCCACCGCCCTCCCAAGGGGCTCGGGAGCTCAGGTGCCCGCAGTTCAATCTCC 1450
Db 910 CTCGCCACCGCCCTCCCAAGGGGCTCGGGASCTCAGGTGSSCCAGTTCAATCTCC 969
QY 1451 GCTGCTGCTGCTGGCCCTTACCTTCCCAAGCGCTCCAGTCTCTGGCAGTTTCTGGAAATG 1510
Db 970 GCTGCTGCTGCTGGCCCTTACCTTCCCAAGCGTCCAGTCTCTGGCAGTTTGGAAATG 1029
QY 1511 AAGGGTTCTGGCTGCCCAACTCTGTAAGGGCAGAGGTGGAGGGTGGGGGGCGCTGAGT 1570
Db 1030 AAGGGTTCTGGCTGCCCAACTCTGTAAGGGCAGAGGTGGAGGGTGGGGGGCGCTGAGT 1089
QY 1571 AAGGACTCAGGGGCAATGCTGCCCTCATCTCATTTCAAAACCCACCTAGTTTCCCTG 1630
Db 1090 AAGGACTCAGGGGCAATGCTGCCCTCATCTCATTTCAAAACCCACCTAGTTTCCCTG 1149
QY 1631 AAGGAAATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCAGGCGGGCGGAATCC 1690

Db 1150 AAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGCGCGGCGGAATCC 1209
QY 1691 CTTCCCTGTCAAGAGTGTCACTTCGGCTGCCCTCTGCTGCTTCTGTGTGTGTGAGCAGAA 1750
Db 1210 CTTCCCTGTCAAGAGTGTCACTTCGGTGGCCCTCTGCTGCTTCTGTGTGTGTGAGCAGAA 1269
QY 1751 GTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTTGCACCTTCCCTGACCCCGTCTAATAT 1810
Db 1270 GTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTTGCACCTTCCCTGACCCCGTCTAATAT 1329
QY 1811 ATAAATATAGAGATGTCTATGGCTG 1837
Db 1330 ATAAATATAGAGATGTCTAARMMWR 1356
RESULT 8
US-10-072-036-38
; Sequence 38, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BURON
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 38
; LENGTH: 1896
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Brk1 fusion construct
; NAME/KEY: CDS
; LOCATION: (1)..(1893)
US-10-072-036-38

Query Match 62.0%; Score 1138.4; DB 5; Length 1896;
Best Local Similarity 99.9%; Pred. No. 2.6e-308;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 ATGGCGGCGCGCGGCTCAGGGGGCGGGGCGGGAGCCCGTAGAAACCGAGGGGGTC 71
Db 757 ATGGCGGCGCGCGGCTCAGGGGGCGGGGCGGGAGCCCGTAGAACCGAGGGGGTC 816
QY 72 GGGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTGACGTGGGCGCG 131
Db 817 GGGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTGACGTGGGCGCG 876
QY 132 CGCTACACGAGTTGCAGTACATCGGCGAGGGGCGGTACGGCATGCTCAGCTCGGCCTAT 191
Db 877 CGCTACACGAGTTGCAGTACATCGGCGAGGGGCGGTACGGCATGCTCAGCTCGGCCTAT 936
QY 192 GACCACTGCGCAAGACTCGCTGGCCATCAAGAGATCAGGCCCTTCCGAACATCAGACC 251
Db 937 GACCACTGCGCAAGACTCGCTGGCCATCAAGAGATCAGGCCCTTCCGAACATCAGACC 996
QY 252 TACTGCCAGCGACGCTCCGGGAGATCCAGATCTCTGCTCGCTTCGCCCATGAGAAATGTC 311
Db 997 TACTGCCAGCGACGCTCCGGGAGATCCAGATCTCTGCTCGCTTCGCCCATGAGAAATGTC 1056
QY 312 ATCGGCATCCGAGACATTTCTGGGGGCTCCACCTGGAGCCATGAGAGATGCTACATT 371
Db 1057 ATCGGCATCCGAGACATTTCTGGGGGCTCCACCTGGAGCCATGAGAGATGCTACATT 1116

372 GTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAGCCGAGCGTGGCAAT 431
Db GTGCAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAGCCGAGCGTGGCAAT 1176
432 GACCATATCTGCTACTTCCCTCTACAGATCTCGGGGCTCAAGTACATCCACTCGGC 491
Db GACCATATCTGCTACTTCCCTCTACAGATCTCGGGGCTCAAGTACATCCACTCGGC 1236
492 AACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACCTCGACCTT 551
Db AACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCAGCAACACCACCTCGACCTT 1296
552 AAGATTGATTTGGGCTGGCCCGGATTCGCGATTCCTGAGCATGACACACCGGCTTC 611
Db AAGATTGATTTGGGCTGGCCCGGATTCGCGATTCCTGAGCATGACACACCGGCTTC 1356
612 CTGACGAGTATGGCTACGCGCTGTACCGGCTGACCGGCTCCAGATCATGCTGAATCCAA 671
Db CTGACGAGTATGGCTACGCGCTGTACCGGCTGACCGGCTCCAGATCATGCTGAATCCAA 1416
672 GGCTATACCAAGTCCATCGACATCTGCTCTGTGGGCTGCAATCTGGCTGAGATGCTCTCT 731
Db GGCTATACCAAGTCCATCGACATCTGCTCTGTGGGCTGCAATCTGGCTGAGATGCTCTCT 1476
732 AACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATCTTGGGATC 791
Db AACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATCTTGGGATC 1536
792 CTGGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAGGCCGGAATCTAC 851
Db CTGGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAGGCCGGAATCTAC 1596
852 CTACAGTCTCTGCCCTCCAGACCAAGTGCTGTGGGCTCAAGCTTTTCCCAAGTCAAG 911
Db CTACAGTCTCTGCCCTCCAGACCAAGTGCTGTGGGCTCAAGCTTTTCCCAAGTCAAG 1656
912 TCCAAAGCCCTTGACCTGTGGACCGGATGTAACCTTTAAACCCCAATAAAGGATCACA 971
Db TCCAAAGCCCTTGACCTGTGGACCGGATGTAACCTTTAAACCCCAATAAAGGATCACA 1716
972 GTGAGGAAGCGCTGGCTCAACCTTACCTGGAGGAGTACTATGACCGGAGGATGAGCCA 1031
Db GTGAGGAAGCGCTGGCTCAACCTTACCTGGAGGAGTACTATGACCGGAGGATGAGCCA 1776
1032 GTGSCCGAGGAGCCCTTCACTTCGCGATGGAGCTGGATGACCTTACCTAAAGGAGCGCTG 1091
Db GTGSCCGAGGAGCCCTTCACTTCGCGATGGAGCTGGATGACCTTACCTAAAGGAGCGCTG 1836
1092 AAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGGAGGCCCTTAG 1151
Db AAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGGAGGCCCTTAG 1896

RESULT 9

US-09-925-297-323
; Sequence 323, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR APPLICATION NUMBER: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 323
; LENGTH: 1214
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1203)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-925-297-323
Query Match 58.6%; Score 1076; DB 3; Length 1214;
Best Local Similarity 95.7%; Pred. No. 7.1e-291;
Matches 1137; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
644 GGCCCCAGAGATCATCTGAACTCCAAAGGCTATACAAAGTCCATCGACATCTGGTCTGT 703
Db GGCCCCAGAGATCATCTGAACTCCAAAGGCTATACAAAGTCCATCGACATCTGGTCTGT 87
704 GGCTGCAATTCGGCTGAGATGCTCTAAACCGGCCCATCTTCCCTGGGAAAGCACTACCT 763
Db GGCTGCAATTCGGCTGAGATGCTCTAAACCGGCCCATCTTCCCTGGGAAAGCACTACCT 147
764 GGATCAGCTCAACCAATTCGGGCTCCGGCTCCCATCCAGGAGGACCTGAATTC 823
Db GGATCAGCTCAACCAATTCGGGCTCCGGCTCCCATCCAGGAGGACCTGAATTC 207
824 TATCATCAACATGAAGGCCGGAATACCTACAGTCTCTGCCCCTCCAAGACCAAGTGGC 883
Db TATCATCAACATGAAGGCCGGAATACCTACAGTCTCTGCCCCTCCAAGACCAAGTGGC 267
884 TTGGGCCAAAGCTTTTCCCAAGTCAAGTCTCAAAAGCCCTTGACCTGCTGGACCGGATGT 943
Db TTGGGCCAAAGCTTTTCCCAAGTCAAGTCTCAAAAGCCCTTGACCTGCTGGACCGGATGT 327
944 AACCTTTAACCCCAATAAAGGATCAAGTGGAGGAGGCTGGCTCACCCTACCTGGA 1003
Db AACCTTTAACCCCAATAAAGGATCAAGTGGAGGAGGCTGGCTCACCCTACCTGGA 387
1004 GCAGTACTATGACCGGAGTCAAGTGGGCGGAGGAGCCCTTCACTTTCGCCATGGA 1063
Db GCAGTACTATGACCGGAGTCAAGTGGGCGGAGGAGCCCTTCACTTTCGCCATGGA 447
1064 GCTGGATGACCTTAAAGGAGCGGCTGAAGAGCTCATCTTCAAGGAGACAGACGCTT 1123
Db GCTGGATGACCTTAAAGGAGCGGCTGAAGAGCTCATCTTCAAGGAGACAGACGCTT 507
1124 CAGCCCGGAGTCTGGAGGCCCTTAGCCACAGACATCTTGCACCCCTGGGCGCTG 1183
Db CAGCCCGGAGTCTGGAGGCCCTTAGCCACAGACATCTTGCACCCCTGGGCGCTG - 566
1184 GAACAGACTGGCAAGAGGCAAGAGTCACTGAGGCGCTCTGTCAACCCAGGACCTGCCT 1243
Db -----GGACCTGCCT 576
1244 CTTGCTTGCCTTCTCCCGCAGACTGTTAGAAAAATGGAACACTGTGCCAGCCCGGACCT 1303
Db CTTGCTTGCCTTCTCCCGCAGACTGTTAGAAAAATGGAACACTGTGCCAGCCCGGACCT 636
1304 TGGCAGCCAGGCGGGGTGGAGATGGGCTGGGCACTCTCTCTCTTTGCTGAGGCTC 1363
Db TGGCAGCCAGGCGGGGTGGAGATGGGCTGGGCACTCTCTCTCTTTGCTGAGGCTC 696
1364 CAGCTTCAGGAGGCAAGGCTTCTCTCCCAACCCGCTCCCAAGGAGGCTCGGGA 1423
Db CAGCTTCAGGAGGCAAGGCTTCTCTCCCAACCCGCTCCCAAGGAGGCTCGGGA 756
1424 GCTCAGTGGCCCGCAGTTCAATCTCCCGCTGCTGCTGCGGCTTACCTTCCCAAGCG 1483
Db GCTCAGTGGCCCGCAGTTCAATCTCCCGCTGCTGCTGCGGCTTACCTTCCCAAGCG 816
1484 TCCAGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGGCTGCGGCCAACTGCTGGAAGGC 1543
Db TCCAGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGGCTGCGGCCAACTGCTGGAAGGC 876
1544 AGAGTGGAGGTTGGGGGCTGAGTGAAGGATCAAGGCGGCTGCTGCGGCCCTCATCT 1603
Db AGAGTGGAGGTTGGGGGCTGAGTGAAGGATCAAGGCGGCTGCTGCGGCCCTCATCT 936

QY 1604 CATTCAAAACCCACCTAGTTTCCCTGAAGGAACATTCTTCTAGTCTCAAGGGCTAGCATC 1663
Db 937 CATTCAAAACCCACCTAGTTTCCCTGAAGGAACATTCTTCTAGTCTCAAGGGCTAGCATC 996
QY 1664 CTTGAGGAGCAGGCGCGGCGGAATCCCTCCCTGTCAAAGCTGTCACTTCGCGTGCCT 1723
Db 997 CTTGAGGAGCAGGCGCGGCGGAATCCCTCCCTGTCAAAGCTGTCACTTCGCGTGCCT 1056
QY 1724 CGTGTCTTCTGTGTGTGAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGGGGCC 1783
Db 1057 CGTGTCTTCTGTGTGTGAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGGGGCC 1116
QY 1784 CTGCCACCTCCCTGACCGCTCAATATATAATATAGAGATGTGTCTA 1831
Db 1117 CTGCCACCTCCCTGACCGCTCAATATATAATATAGAGATGTGTCTA 1164

RESULT 10

US-10-114-270-109
; Sequence 109, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Baha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Gasman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 109
; LENGTH: 1147
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1078)
US-10-114-270-109
Query Match 56.8%; Score 1043.2; DB 7; Length 1147;
Best Local Similarity 94.7%; Pred. No. 11e-281;
Matches 1115; Conservative 0; Mismatches 3; Indels 60; Gaps 1;
QY 12 ATGGCGGCGGCGGCGGCTCAGGGGGGCGGGGAGCCCGTAGAACCGAGGGGTC 71
Db 1 ATGGCGGCGGCGGCGGCTCAGGGGGGCGGGGAGCCCGTAGAACCGAGGGGTC 60
QY 72 GCGCGGGGGTCCCGGGGAGGTGAGATGTGTGAAGGGGAGCCGTTGACGCTGGGCCG 131
Db 61 GCGCGGGGGTCCCGGGGAGGTGAGATGTGTGAAGGGGAGCCGTTGACGCTGGGCCG 120
QY 132 CGCTACACGCAGTTCCAGTACATCGCGGAGGGCGGTACGGCATGTGTGAGCTCGGCCTAT 191
Db 121 CGCTACACGCAGTTCCAGTACATCGCGGAGGGCGGTACGGCATGTGTGAGCTCGGCCTAT 180
QY 192 GACCAGTGGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCTTGAACATCAGACC 251
Db 181 GACCAGTGGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCTTGAACATCAGACC 240
QY 252 TACTGCCAGGCGACGCTCCGGGAGATCCAGATTCCTGTGCGCTTCGGCATGAGATGTC 311
Db 241 TACTGCCAGGCGACGCTCCGGGAGATCCAGATTCCTGTGCGCTTCGGCATGAGATGTC 300
QY 312 ATCGGCATCCGAGACATTTCTGGGGGCTCCACCTGGAAGCCATGAGAGATGTCATATT 371
Db 301 ATCGGCATCCGAGACATTTCTGGGGGCTCCACCTGGAAGCCATGAGAGATGTCATATT 360
QY 372 GTGCAGGAGCTCATGAGAGACTGACCTGTACAAAGTTGCTGAAAGGCGAGCTGAGCAAT 431
Db 361 GTGCAGGAGCTCATGAGAGACTGACCTGTACAAAGTTGCTGAAAGGCGAGCTGAGCAAT 420
QY 432 GACCATATCTGCTACTTCTTACAGATTCCTGGGGGCTCAAGTACATCCATCCGCC 491
Db 421 GACCATATCTGCTACTTCTTACAGATTCCTGGGGGCTCAAGTACATCCATCCGCC 480
QY 492 AACGTGCTCCAGGAGATCTAAAGCCCTCAACCTGCTCATCAACACACCTCGGACCTT 551
Db 481 AACGTGCTCCAGGAGATCTAAAGCCCTCAACCTGCTCATCAACACACCTCGGACCTT 540
QY 552 AAGATTTGTGATTTCCGGCTGCGCGGATTGCGGATTCCTGAGCATGACACACCGGCTTC 611
Db 541 AAGATTTGTGATTTCCGGCTGCGCGGATTGCGGATTCCTGAGCATGACACACCGGCTTC 600
QY 612 CTGACGGAGTATGTGGCTACCGCTGGTACCGGGGCCCGAGAGATCATGCTGAATCCCAAG 671
Db 601 CTGACGGAGTATGTGGCTACCGCTGGTACCGGGGCCCGAGAGATCATGCTGAATCCCAAG 660
QY 672 GGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCTATCTGGCTGAGATGCTCTCT 731
Db 661 GGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCTATCTGGCTGAGATGCTCTCT 720
QY 732 AACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACATTTCTGGGCATC 791
Db 721 AACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACATTTCTGGGCATC 780
QY 792 CTGGGCTCCCATCCAGGAGACCTGAATTGTATCATCAATGAGGCCCGAACTAC 851
Db 781 CTGGGCTCCCATCCAGGAGACCTGAATTGTATCATCAATGAGGCCCGAACTAC 840


```
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-171-311-126

Query Match 35.2%; Score 646.6; DB 5; Length 1611;
Best Local Similarity 73.2%; Pred. No. 1.6e-170;
Matches 829; Conservative 0; Mismatches 304; Indels 0; Gaps 0;

QY 1 GAGGATGAGAGTGGCGGCGCGCGCTCAGGGGCGCGGGGCGGGGAGCCCCGTAGAA 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CCGAGGGGTCGCGCGCGGGTCCCGGGGAGGTGAGATGGTGAAGGGCAGCGTTCG 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 326 ACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 385
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ACGTGGCGCGCGCTACAGCGAGTTCGAGTACATCGCGCGAGGCGCGGTACGGCATGCTCA 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 386 ACGTGGCGCGCGCTACACCACTCTCGTACATCGCGCGAGGGCGCTACGGCATGTCT 445
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GCTGGCGCTATGACCAAGTGGCGGCAAGACTCGCGTGGCGCATCAAGAAAGATCAGCCCCCTTCG 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 446 GCTCTGTTATGATAATGTCAACAAAGTTCGAGTAGCTATCAAGAAATCAGCCCCCTTTG 505
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 AACATCAGACCTACTCCAGCGCAGCTCCGGGAGATCCAGATCCTGCTGCGCTCCGCC 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 506 AGCACCAGACCTACTCCAGAGAACCTGAGGGAGATAAAATCTTACTGCGCTTCAGAC 565
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 ATCAGAAATGTCATCGGCATCCGAGACATTCGCGGCGTCCACCTCGGAAGCCATGAGAG 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 566 ATGAGAACATCATGTGAATCAATGACATATTTCAGACACCAACCATCGAGCAATGAAG 625
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 ATGTCTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAGCCAGC 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 626 ATGTATATATAGTACAGGACCTCATGGAACAGATCTTTTACAGCTCTTGAAGACACAAC 685
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 AGCTGAGCAATGACCATATCTGTACTCTCTACAGATCCCTGCGGGGCTCAAGTACA 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 686 ACCTCAGCAATGACCATATCTGTATTTTCTTACCAGATCTTCAGAGGGTTAAATATA 745
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 TCCACTCCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCACCTGCTCATCAACACCA 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 746 TCCATTTCAGCTAAGTTCTGCACCGTGACCTCAAGCTTCCACCTGCTGCTCAACACA 805
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 CTTGCGACCTTAAGATTGTGATTTTCGGCTGCGCGGATGTCGGATCTCTGAGCATGACC 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 806 CTTGTGATCTCAAGATCTGTGACTTTTGGCCCTGGCCGCTGTTCAGATCCAGACCATGATC 665
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

601 ACACCGGCTTCTGACGGAGTATGTGGCTAGCGCTGGTACCGGGCCCGAGAGATCATGC 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
866 ACACAGGGTTCTGACAGAAATATGTGGCCACACGTTGGTACAGGGCTCCAGAAATATGT 925
QY 661 TGAATCCAAAGGGCTATACCAAGTCCATCGATCCATCTGGTCTGTGGGCTGCACTTGGCTG 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
926 TGAATTCCAAAGGGCTACACCAAGTCCATTTGATTTGGTCTGTGTAGGCTGCATTTCTGGCAG 985
QY 721 AGATGCTCTTAACCGGGCCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCCACA 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
986 AAATGCTTTTCCAAAGGGCCCATCTTTTCCAGGGAAGCATTTATCTTGACGAGCTGAATCACA 1045
QY 781 TTCTGGGATCTGGGCTCTCCCATCCAGAGGAGCTGAAATGTATCATCAACATGAAGG 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1046 TTTTGGGTATTTCTTGATCCCATCATCAAGAAGAGCTGAAATGTATATAAATTTAAAG 1105
QY 841 CCCGAAACTTACCTACAGTCTCTGCGCTCAAGACCAAGTGGCTTTGGGCAAGCTTTTCC 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1106 CTAGGAACTATTGTGCTTCTCTCCACACAAATAAAGTGCATGGAACAGGCTGTTC 1165
QY 901 CCAAGTCAGACTCCAAAGCCCTTGACCTGGACCGGATGTTAACCTTTAACCCCAATA 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1166 CAAATGCTGACTCCAAAGCTCTGGACTTATTGGGACAAATGTTGACATTTCAACCCACACA 1225
QY 961 AACGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACTGGAGCAGTACTATGACCCGA 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1226 AGAGGATTCAGTAGACAGGCTCTGGCCACCATATCTGGAGCAGTATTACGACCCGA 1285
QY 1021 CGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTCCGCAATGGAGCTGGATGACCTTACCTTA 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1286 GTGACGAGCCCATCGCGAAGCACCATTCAAGTTCCGACATGGAATTTGGATGACTTGCCTTA 1345
QY 1081 AGGAGCGGCTGAGAGAGTCACTTCTCCAGGAGACAGCAGCTTCCAGCCCGGA 1133
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1346 AGGAAAGCTAAAGAACTAATTTTGAAGAGACTGTAGATTTCCAGCCCGGA 1398

RESULT 13
US-10-301-822-114
; Sequence 114, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (328)...(1410)
; US-10-301-822-114
```

Query Match 35.2%; Score 646.6; DB 6; Length 1611;
Best Local Similarity 73.2%; Pred. No. 1.6e-170;
Matches 829; Conservative 0; Mismatches 374; Indels 0; Gaps 0;

Qy 1 GAGGAGTGGAGATGGCGCGCGCGCTCAGCGGGGCGGGGCGGGGAGCCCGTAGAA 60
Db 266 GAGCGCGCGCGCGCGCGCTCAGCGGGGCGGGGCGGGGAGCCCGTAGAA 325

Qy 61 CCGAGGGGGTCCGCGCGCGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGCCGTTCG 120
Db 326 ACATGGCGCGCGCGCGCGCGCGCGCGCGCGCGAGATGGTCCGCGGGCAGGTTCG 385

Qy 121 ACGTGGCGCGCGGTACACGAGTTCAGTACATCGCGGAGGGCGGTACGGCATGGTCA 180
Db 386 ACGTGGCGCGCGGTACACCAACCTCTCGTACATCGCGGAGGGCGCGGTACGGCATGGTGT 445

Qy 181 GCTCGGCTATGACCAAGTGGCGGAGACTCGGTGGGCATCAAGAAAGATCAGCGCCCTTCG 240
Db 446 GCTCTGCTTATGATATGTCACAAAGTTTCAGTACGTATCAAGAAATCAGCCCTTTG 505

Qy 241 AACATCAGACTACTGCGCAGCGACGCTCCGGGAGATCCAGATCTCTGCGCTTCGCC 300
Db 506 AGCACAGACTACTGCGCAGAACCTCGAGGAGATAAATCTTACTCGGCTTCAGAC 565

Qy 301 ATGAGATGTATCGGCATTCGAGACATTCCTGGGGCGTCCACCCTGGAAGCCATGAGAG 360
Db 566 ATGAGAACATCATTTGAATCAATGACATTTTCAGGACCAACCATCGAGCAATGAAG 625

Qy 361 ATGTCTACATTTGCAGGACTGATGAGACTGACCTGATACAGTTGCTCAAGAGCCAGC 420
Db 626 ATGTATATATAGTACAGGACTCATGGAACAGATCTTTTAAAGCTCTTGAAGACAAAC 685

Qy 421 AGCTGAGCAATGACATATCTGTACTTCTCTACAGATCTCTCGGGGCTCAAGTACA 480
Db 686 ACCTCAGCAATGACATATCTGTATTTCTTACAGATCTCTCAGAGGTTAAATATA 745

Qy 481 TCACCTCCGCAACGTCCTCACCGAGATTAAGCCCTCCAACTGCTCATCAACACCA 540
Db 746 TCCATTGAGTAAAGTTCTGACCGTCAAGCTCAAGCTTCCAACTGCTCATCAACACCA 805

Qy 541 CCTCGACCTTAAGATTTGATTTCCGCTGGCGCGGATGCGGATCTGAGCATGACC 600
Db 806 CCTGTGATCTCAAGATCTGTACTTTGGCTGGCGCGGTGTCAGATCCAGCATGATC 865

Qy 601 ACACCGCTTCTCAGCGAGTATGCTACGCGCTGTGTACCGGGCCCGAGATCATGC 660
Db 866 ACACAGGTTCTCAGCAGATATGTCGCCACACGTTGTCAGGGCTCCAGAAATATGT 925

Qy 661 TGAATCCAAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCAATCTGGCTG 720
Db 926 TGAATCCAAAGGGCTACACCAAGTCCATTTGATATTTGGTCTGTAGGCTGCATTCGGCAG 985

Qy 721 AGATGCTCTTAACCGGGCCATCTTCCCTGGCAGCACTACTGTGATCAGCTCAACACA 780
Db 986 AAATGCTTTCAACAGGGCCATCTTTCCAGGGAAGCAATTTTTCAGCAGCTGAATCACA 1045

Qy 781 TTCTGGGCATCTCGGCTCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAAG 840
Db 1046 TTTTGGGGTATTTCTTGGATTCCTCATCAAGAGACCTGAATTTGATTAATTAATTAAG 1105

Qy 841 CCCGAAACTACTACAGTCTCTGCCCTCCAAAGACCAAGGTGGCTTGGGCCAAGCTTTTC 900
Db 1106 CTAGGAATATTTGCTTCTTCCACACAAAATAGGTGCCATGGAACAGGCTGTTC 1165

Qy 901 CCAAGTCAGATCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAACCCCAATA 960
Db 1166 CAATGCTGATCCAAAGCTCTGGAATTTATTTGGAACAAAATTTTGACATTTCAACCCACA 1225

Qy 961 AACGATCAGATGGAGAGCGCTGCTCAACCTTACCTGAGCAGTACTATCAGCCGA 1020
Db 1226 AGAGATGAAGTGAACAGGCTCTGCGCCACCCATATCTGGAGCAGTATTCAGCCCGA 1285

Qy 1021 CGGATGACCAAGTGGCGGAGGAGCCCTTCACTTTCGCCATGGAGCTGGATGACCTACCTA 1080

Db 1286 GTGACAGGCCATCGCCGAGCACCAATTCAGTTTCGACATGGATTCGATTCGCTTA 1345
Qy 1081 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCGGA 1133
Db 1346 AGGAAAAGCTAAAAGAACTAAATTTTGAAGAGACTGCTAGATTCAGCCAGGA 1398

RESULT 14
US-10-305-720-1263
; Sequence 1263; Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1263
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g23878
US-10-305-720-1263

Query Match 34.9%; Score 642; DB 6; Length 2791;
Best Local Similarity 74.9%; Pred. No. 3.6e-169;
Matches 804; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

Qy 60 ACCGAGGGGTCCGCGCGGGGTCCCGGGGAGGTGGAGATGTTGAAGGGGCGAGCCGCTC 119
Db 96 ACCTGTGATGGCGCGCGCGCGCGCGCGCGCGCGCGAGATGTTCCGCGGCGAGGTGTC 155

Qy 120 GAGTGGCGCGCGCTACACGAGTTTCAGTATCATCGCGAGGGCGCGGTACGGCATGGTC 179
Db 156 GAGTGGCGCGCGCTACACCAACCTCTCGTACATCGCGAGGGCGCGCTACGGCATGGTG 215

Qy 180 AGCTGGCGCTATGACACGTCGCGAGATCTCGGTGGCCATCAAGAGATCAGCCCCCTC 239
Db 216 TGCTCTGCTTATGATATGTTCAACAAAGTTTCGAGTAGCTATCAAGAAATCAGCCCCCTT 275

Qy 240 GAACATCAGACCTTACGCCAGCGCACGCTCCCGGAGATCCAGATCTCTGCTGCGCTCCGC 299
Db 276 GAGCACAGACCTTACTGCCAGAGAACCTCGAGGAGATATAAATCTTACTGCGCTTCAGA 335

Qy 300 CATGAGAAATGTCATCGGCATCCGAGACATTTCTCGGGCGTCCACCTTGGAGGCCATGAGA 359
Db 336 CATGAGAACATCATTTGGAATCATGACATTTATTCGAGACCAACCATCGAGCAATGAAA 395

Qy 360 GATGCTTACATTTGTCAGGACCTGATGAGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 419
Db 396 GATGTTATATAGTACAGGACCTCATGGAACAGATCTTTTCAAGCTCTTTGAAGACAAA 455

Qy 420 CAGTGGAGCAATGACCATATCTGCTACTTCTCTTACAGATCTCTGCGGGCGCTCAAGTAC 479
Db 456 CAGCTCAGCAATGACCATATCTGCTATTTCTTCTACAGATCTCTCAGAGGCTTAAATAT 515

Qy 480 ATCCACTTCGCCCAAGCTGCTCCACCGAGATCTAAAGCCCTTCCAACTGCTCATCAACACC 539
Db 516 ATCATTCAGCTAACGTTCTGCAACCGTGAATTTTCAAGCTTCCAACTGCTCATCAACACC 575

Qy 540 ACCTGCGACCTTAAGATTTGTGATTTTCGGCTCGCGCGGATTCGCGATCTCTGAGCATGAC 599
Db 576 ACCTGTGATCTCAAGATCTGTGACTTTGGCGCTGGCGCGTGTGAGATCCAGACCATGAT 635

Qy 600 CACACCGGCTTCTGAGCGGAGTATGTGCTACGGCTGTTACCGGGCCCGCAGAGATCATG 659

```

Db      636  CACACAGGGTTCCTGACAGAAATATGTGGCCACACAGTTGGTACAGGGCTCCAGAAATTATG 695
QY      660  CTGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTCTGGCT 719
Db      696  TTGAATTCGAAGGGCTACCAAGTCCATGATTAATTGGTCTGTGGCTGCATTCTGGCA 755
QY      720  GAGATGCTCTTAACCGGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACAC 779
Db      756  GAATGCTTCTTAACAGGGCCATCTTCCAGGGAGCATTAATCTGACAGCTGAACAC 815
QY      780  ATTCTGGGCATCTGGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAG 839
Db      816  ATTTTGGGTATTCTTGGATCCCATCACAAGAGACCTGAATGTATATAATTAATTTAAA 875
QY      840  GCCGGAATCTACCTACAGTCTCTGCCCTCCAGACCAAGGTGCTTGGGCCAAGCTTTTC 899
Db      876  GCTAGGAATTAATTTGCTTCTCTCCACACAAAAATAAGGTGCCATGGAACAGGGCTGTC 935
QY      900  CCCAAGTTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACCTTTAACCCCAAT 959
Db      936  CCAATGCTGACTCCAAAGCTCTGGACTTATTTGGACAAAATGTTGACATTTCAACCCAC 995
QY      960  AAACGGATCACAAGTGGAGGAGGGCTGGCTACCCCTACCTGAGGAGTACTATGACCCG 1019
Db      996  AAGAGGATTGAAGTAGAACACAGGCTCTGGGCCACCCATATCTGGAGCAGTATTACGACCC 1055
QY      1020  ACGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTTACCT 1079
Db      1056  AGTGACGAGCCCATCGCCGAGACCAATTTCAAGTTGACATGGAATTTGATGATGCTGCT 1115
QY      1080  AAGGAGCGGCTGAAGGAGCTCATCTTTCCAGGAGACAGCAACGCTTCCAGCCCGGA 1133
Db      1116  AAGGAAAAGCTCAAGAACTAATTTTGAAGAGACTGCTAGATTCAGCCAGGA 1169

RESULT 15
US-10-756-149-513
; Sequence 513, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 513
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-513

Query Match      34.9%; Score 642; DB 9; Length 2791;
Best Local Similarity 74.9%; Pred. No. 3.6e-169;
Matches 804; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

QY      60  ACCGAGGGGTCCGGCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGACCGCTTC 119
Db      96  ACCTGTGATGGCCGGCGGGCGGGCGGAGATGGTCCGGGCGAGGTGTC 155
QY      120  GACGTGGGCGCGCTACACGAGTTGCGAGTACATCGGCGAGGGCGGCTACGGCATGGTC 179
Db      156  GACGTGGGCGCGGCTACACCAACCTCTCGTACATCGGCGAGGGCGCTACGGCATGGT 215
QY      180  AGCTCGGCTATCACACGTCGCCAGACTCGGTCGCCATCAAGAGATCAGCCCTTC 239
Db      216  TGTCTGCTTATGATATGTCACAAAGTTTCAGTAGTACATCAAGAAATCAGCCCTTC 275
QY      240  GAACATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCCTGCTGGCTTCGCG 299

```

```

Db      276  GAGCACCAAGACCTACTGCGCAGAGAAACCCCTGAGGGAGATAAAAAATCTTACTGGCTTCAGA 335
QY      300  CATGAGATGTCTATCGGGATCCGAGACATTTCTGCGGGCTCCACCTCGGAAGCCATGAGA 359
Db      336  CATGAGAACATCATTTGGAAATCAATGACATTAATTCGAGACCAACCAATCGAGGAATGAAA 395
QY      360  GATGCTCTCATTTGTGCGAGGACCTGATGGAGACTGACCTGTGTACAAGTTGCTGAAAAAGCCAG 419
Db      396  GATGTATATATAGTACAGGACCTCATGGAACAGATCTTTACAAGCTCTTGAAGACACAA 455
QY      420  CAGCTGAGCAATGACATATCTGCTACTCTTACAGGATCTGCGGGGCTCTAAGTAC 479
Db      456  CACCTCAGCAATGACATATCTGCTATTTCTCTACAGATCTCTCAGAGGGTTAAAAAT 515
QY      480  ATCCACTCGGCAACGTGCTCCACAGATCTTAAAGCCCTCCAACTGCTCATCAACACC 539
Db      516  ATCCATTCAGCTTAAAGCTTCTGCACCGTACCTCAAGCCCTTCCAACTGCTCTCAACACC 575
QY      540  ACTCTGCGACTTAAAGATTTGTGATTTGCGCCCTGGGCCCGGATTCGCCATCTCTGAGCATGAC 599
Db      576  ACTCTGATCTCAAGATCTGTGACTTTGGCCCTGGCCCGTGTTCAGATCCAGACCATGAT 635
QY      600  CACACGGGCTTCTGACGAGTATGTGCTACGCGTGTGTACGCGGCCCCACAGATCATG 659
Db      636  CACACAGGGTTCCTGACAGAAATATGTGGCCACACGTTGGTACAGGGCTCCAGAAATTTATG 695
QY      660  CTGAATCTCAAGCGGCTATATCAAGTCCATCGACATCTGTGCTGTGGGCTGCTTCTGGCT 719
Db      696  TTGAATTCGAAGGGCTACCAAGTCCATTTGATATTTGGTCTGTAGGCTGCAATCTGGCA 755
QY      720  GAGATGCTCTTAACCGGCCATCTTCCCTGGCAAGCACTACTCGATCAGCTCAACCCAC 779
Db      756  GAAATGCTTTCTAAACAGGCCATCTTCCAGGGAAGCATTAATCTTGACAGCTGAACCCAC 815
QY      780  ATTCTGGGATCCTGGGCTCCCATCCAGGAGGACCTGGAATTTGTATCATCAACATGAAG 839
Db      816  ATTTTGGGTATTCTTGGATCCCATCAAGAGAGACCTGAATTTGTATATAATTTAAAA 875
QY      840  GCCGGAATCTACCTACAGTCTCTGCCCTCCAGACCAAGGTGCTTGGGCCAAGCTTTTC 899
Db      876  GCTAGGAATTAATTTGCTTCTCTTCCACAAAAATAAGGTGCCATGGAACAGGGCTGTTTC 935
QY      900  CCAAGTCAAGCTCCAAAGCCCTTGAACCTGCTGGACCGGATGTTAACCTTTAACCCCAAT 959
Db      936  CCAATGCTGACTCCAAAGCTCTGGACTTATTTGGACAAAATGTTGACATTTCAACCCAC 995
QY      960  AAACGGATCACAAGTGGAGGAGGGCTGGCTACCCCTACCTGAGGAGTACTATGACCCG 1019
Db      996  AAGAGGATTGAAGTAGAACACAGGCTCTGGGCCACCCATATCTGGAGCAGTATTACGACCC 1055
QY      1020  ACGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTTACCT 1079
Db      1056  AGTGACGAGCCCATCGCCGAGACCAATTTCAAGTTGACATGGAATTTGATGATGCTGCT 1115
QY      1080  AAGGAGCGGCTGAAGGAGCTCATCTTTCCAGGAGACAGCAACGCTTCCAGCCCGGA 1133
Db      1116  AAGGAAAAGCTCAAGAACTAATTTTGAAGAGACTGCTAGATTCAGCCAGGA 1169

```

Search completed: February 7, 2006, 15:20:13
Job time : 1587.19 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1708.4	93.0	1866	7	US-10-955-054A-5		Sequence 5, Appli
2	1059.8	57.7	1833	8	US-11-136-527-284		Sequence 284, App
3	646.6	35.2	1611	8	US-11-186-284-114		Sequence 114, App
4	630.8	34.3	1725	8	US-11-136-527-1988		Sequence 1988, Ap
5	275	15.0	1577	8	US-11-136-527-2485		Sequence 2485, Ap
6	273.4	14.9	3132	8	US-11-136-527-2850		Sequence 2850, Ap
7	273.2	14.9	1888	7	US-10-770-726-25		Sequence 25, Appl
8	260.8	14.2	2737	7	US-10-909-125-861		Sequence 861, App
9	233	12.7	1665	8	US-11-136-527-3828		Sequence 3828, Ap
10	225	12.2	1400	8	US-11-136-527-7934		Sequence 7934, Ap
11	203.8	11.1	3824	8	US-11-136-527-2977		Sequence 2977, Ap
12	166.4	9.1	822	8	US-11-136-527-3508		Sequence 3508, Ap
13	165.8	9.0	1980	8	US-11-136-527-238		Sequence 238, App
14	165	9.0	600	8	US-11-136-527-7604		Sequence 7604, Ap
15	164.2	8.9	600	8	US-11-136-527-4380		Sequence 4380, Ap
16	147.8	8.0	900	8	US-11-136-527-55		Sequence 55, Appl
17	141.8	7.7	3131	7	US-10-770-726-28		Sequence 28, Appl
18	136.2	7.4	2765	8	US-11-136-527-126		Sequence 126, App
19	135.8	7.4	2804	8	US-11-136-527-282		Sequence 282, App
20	135.4	7.4	2211	8	US-11-127-817-10		Sequence 10, Appl
21	135.4	7.4	2372	7	US-10-957-780-11		Sequence 11, Appl
22	135.4	7.4	2372	8	US-11-127-817-9		Sequence 9, Appli

182 CTGGGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 241
198 CTGAGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 257
242 ACATCAGACCTACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 301
258 GCATCAAACTACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 317
302 TGAGAACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 361
318 TGAGAACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 377
362 TGCTCACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 421
378 TGCTCACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 437
422 GCTGAGCAATGACCACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 481
438 GCTGAGCAATGACCACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 497
482 CCACTCGGCAAGCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 541
498 AACTCGGCAAGCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 557
542 CTGCGCACTTAAAGATTTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 601
558 CTGCGCACTTAAAGATTTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 617
602 CACGGGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 661
618 CACTGGGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 677
662 GAATCCAGGGCTTACCAAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 721
678 TAATCCAGGGCTTACCAAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 737
722 GATGCTCTTAAAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 781
738 GATGCTCTTAAAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 797
782 TCTGGGCACTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 841
798 TCTAGGTATATCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 857
842 CCGAACTTACTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 901
858 CCGAACTTACTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 917
902 CAACTGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 961
918 CAACTGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 977
962 ACGGATCAGCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1021
978 GCGCATCAGCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1037
1022 GGATGAGCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1081
1038 AGATGAGCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1097
1082 GGAGCGGCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1141
1098 GGAGCGGCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1157
1142 GGCGGCTTACGAGCAAGACTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1201
1158 GGCGGCTTACGAGCAAGACTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1200
1202 GGCAAGAGCTTCTGAGGAGCTTCTGAGCTTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1261
1201 -----TGTGCTTACCTGCTCTCTC

1262 GCAGACTTTAGAAAATGGAACACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1321
1223 TGAGAACTGTTAGAAAATG--AACTTTGCTCAACCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1279
1322 TGAGAACTGCGGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1381
1280 CCAAGGGTGGGCTGGAACCCCTCTCACTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1337
1382 GGCCTTCTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1441
1338 CTGCCACTCCAGTCCCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1397
1442 CAATCTCCGCTGCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1501
1398 CTGATCTGCTGCTGCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1456
1502 CTGGAATGAAGGGTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1561
1457 CTGGAATGAAGGGTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1510
1562 GGCCTGAGTAGGAGCTCAGGCGCATGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1621
1511 GCACTGAGT-----AGGCTAAGCTCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1562
1622 GTTTCCTGCAAGAACTTCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1681
1563 TTTTCCTGCAAGAACTTCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1622
1682 GCGAATCCCTGCTGCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1740
1623 GCTAACCTTCTGCTGCTGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1681
1741 GTGAGCAGAACTGAGCTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1800
1682 GTGATGAGTGTGAGGCGGCGGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 1741
1801 C----GTCTAATATAAATATAGAGATGCTCTATGG 1834
1742 CTGTATCTAATATAAATATAGAGATGCTCTATGG 1778

RESULT 3

US-11-186-284-114
; Sequence 114, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burt, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 1611
; TYPE: DNA


```
Db 659 TCTGTGACTTTGGCGCTTGGCCGTTTGCAGATCCAGACCAGATCATACAGGGTCTTGA 718
Qy 616 CGAGATGATGGCTTACCGCGCTGTGTACCGGGCCCCAGAGATCATGCTGAACTCCAAGGGCT 675
Db 719 CAGAGTATGTAGCACCGGTTGGTACAGAGCTCCAGAAATATGTTGAATTTCCAAGGGTT 778
Qy 676 ATACCAAGTCCATCGACATCTGTGCTGTGGCTGCGATTTCTGGCTGGAGATGCTCTTAACC 735
Db 779 ATACCAAGTCCATGATATTTGGTCTGTGGGCTGCATCTCTGGCAGAGATGCTATCCAACA 838
Qy 736 GGCCCATCTTCCCTGGCAAGCACTTACCTGGATCAGCTCAACCACTTCTGGGCATCTGG 795
Db 839 GGCCATCTTCTCCAGGAAGCAATTACCTTGACCAAGCTGAATCAATCTGGGATTTCTG 898
Qy 796 GCTCCCATCCAGGAGGACCTGAATTTGTATCATCAACATGAAGGCCGGAACCTACCTAC 855
Db 899 GATCTCCATCACGAAGATCTGAATTTGTATATAAATTTAAAGCTTAGAACTATTTGC 958
Qy 856 AGTCTCTGCCCTCCAAGCAAGGTGCTTGGGCCAAGCTTTTCCCAAGTCAAGCTCCA 915
Db 959 TTTCTCTCCGCACAAAAATAAGGTGCGTGGAAAGAGTTGTTCCCAAACGCTGACTCCA 1018
Qy 916 AAGCCCTTGACCTGCTGGACCGGATGTTAACTTTTACCCCAATAAAGGATCACAGTG 975
Db 1019 AAGCTCTGGATTTACTCGATATAAATGTTGACATTTTAACTTACCTTCAAGAGGATGAAGTG 1078
Qy 976 AGGAAGCGCTGGCTACCCCTTACCTGGAGAGTACTATGACCCGAGTGGAGCCAGTGG 1035
Db 1079 AACAGGCGKRGSCCACCCTGACTCTGGAGCAGTATTTATGACCAAGTGTAGGCCCATTTG 1138
Qy 1036 CCGAGGAGCCCTTCACTCTCGCCATGGAGCTGGATGACCTTACCTTAAGAGCGGCTGAAG 1095
Db 1139 CTGAAGCACCATTCAAGTTTGACATGAGCTGGAGCACTTACCTTAAGGAGAAGCTCAAAG 1198
Qy 1096 AGCTATCTTCCAGGAGACAGCAGCTTCCAGCCCGGA 1133
Db 1199 AACTCATTTTGAAGAGACTGCTCGATTTCCAGCCAGGA 1236

RESULT 5
US-11-136-527-2485
; Sequence 2485, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2485
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2485
Query Match 15.0%; Score 275; DB 8; Length 1577;
Best Local Similarity 56.9%; Pred. No. 9.1e-54;
Matches 597; Conservative 0; Mismatches 425; Indels 27; Gaps 4;

Qy 115 CGTTGAGCTGGGCGCGCTTACAGCAGTTGTGAGTATCATCGGCGAGGGCCGCTACGGCA 174
Db 61 CTGGGAGCTGCCAAGACCTACTTGGCGCGCGCGCACGTGCGCAGCGGGGCTTATGGCG 120
Qy 175 TGGTCAGCTCGGCTTATCACCAGCTGCCAGAGACTCGGCTGCCATCAAGAGATCAGC- 233
Db 121 CGGTGTGCTCGGCCATCAGCAAGGAGCAGGGGAGAGGTGGCCATCAAGAGCTGAGCC 180
Qy 234 --CCCTTGAACATCAGACCTACTGCCAGCGACCGCTCCGGGAGATCCAGATCTCTGTGTC 291
```

```
Db 181 GGCCCTTCCAGTCCGAGATCTTTGCCAAGCGCGCTTACCGGAACTCTCTGCTGTGAAGC 240
Qy 292 GCTTCGCGCCATGAGAAATGTATCGGCATCCGAGA---CATTTGCGGGGGTCCACCGTGG 348
Db 241 ACATGACCATGAGAACGTCATTGGGCTTCTGATGTCTACACACCTGCGCACTTCCGTTTC 300
Qy 349 AAGCCATGAGAGATGTCTACATTTGTGAGGACCTGATGGAGACTGACCTGTACAAAGTTGC 408
Db 301 GAAACTTCCAAGATTTTCTACTCTGTGATGCTTTTTCATGACAGCCGACCTCGAGAAGATAA 360
Qy 409 TGAAGAGCCAGCAGCTGAGCAATGACATCTGCTACTTCTCTACAGATCTCTGGGG 468
Db 361 TG---GGATGAGTTCAGTGAGAGAGAGTCCAGTATTTGGTGTACCAAGTGTCAAAG 417
Qy 469 GCCTCAAGTACATCCACTCTCGCCAAAGTGTCTCAACGAGATCTAAAGCCCTTCAACCTGC 528
Db 418 GTCTAAGTACATCCACTCAGCTGGTATCGTTCAAGGGACCTGAAAGCCAGGCAACCTGG 477
Qy 529 TCATCAACACCACTGCGACCTTAAAGATTTGTGATTTGCGCCCTGGCCCGGATTTGCCGATC 588
Db 478 CCCTGAATGAAGACTGTGAGCTGAAGATCTGAGCTTTGGGCTGGCAGCCGACACGCGATG 537
Qy 589 CTGAGCATGACCACACCGGCTTCTGACGAGATGTGCTACGCGTGTGTACCGGGCTC 648
Db 538 CGAG-----ATGACTGGCTATGTGGTGTGCCGCTGGTACCGGGCTC 579
Qy 649 CAGAGATCATGCTGAACTCCAAGGGCTATACAAAGTCCATCGACATCTGGTCTGTGGGCT 708
Db 580 CTGAGGTGATCTCAGCTGGATGATTAACACAGAGCGGTGGATATCTGCTGTGTGGCT 639
Qy 709 GCATTTGCTGATGATGCTCTTAAACCGGCCATCTTTCCCTGGCAAGCACTACTTGATC 768
Db 640 GCATCATGGCCGAGATGCTTAAACCGGAAAGACACTCTTCAAGGGGTAAAGGACTTACCTG 699
Qy 769 AGCTCAACACACTTCTGGGCTCTCTGGGCTCCCATCCAGAGGAGCTCGAATTTGTATCA 828
Db 700 AGCTGACCCAGATCTCTGAAAGTGAAGTGGGTGCGAGTGGCGAGTTCGTGAGAGCTGA 759
Qy 829 TCACATGAAGGCGCGGAACTACCTACAGTCTCTGCGCTTCCAAGACCAAGGTGGCTTGGG 888
Db 760 AAGCAAGGCGGCGCAATCTACATTCAGTCTCTGCGCCAGAGCGCCCAAGAGGATTTCA 819
Qy 889 CCAAGCTTTTCCCAAGTCAAGTCTCAAGCCCTTGAACCTGCTGGACCGGATGTTACCT 948
Db 820 CACAGCTTTTCCCAAGTCAAGTCTCAAGCTGTAGACCTGCTGGAACAAGTGTGAGC 879
Qy 949 TTAACCCCAATAAAGGATCAGTGGAGGAGCGCTGCGCTCACCCCTACTCTGAGCAGT 1008
Db 880 TGGACGTGGACAAGCGTCTGACCGCTCTCAGGCACTCGCTCACCCCTCTTTGAAACCC 939
Qy 1009 ACTATGACCCGAGGATGAGCGAGTGGCCGAGGAGCCCTTCACTTCCCATGAGAGCTGG 1068
Db 940 TCCGGGACCTTGAGGAGGAGACAGAGGCCCGCAGAGCCATTTGATGATGCTTAGAGCGG 999
Qy 1069 ATGACCTTACCTAAGGAGCGGCTGAAGGAGTCTATTTCCAGGAGACAGAGCTTCCAGC 1128
Db 1000 AGAATCTCAGCGTGGAGCAATGGAACAACACATCTACAAAGAGATCGCCAACTTCAGTC 1059
Qy 1129 CCGGAGTGTGGAGGGCCCCCTAGCCGAGA 1157
Db 1060 CCATAGCCCGAAGGACTCGAGGCGACGA 1088

RESULT 6
US-11-136-527-2850
; Sequence 2850, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
```

; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2850
; LENGTH: 3132
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-11-136-527-2850

Query Match 14.9%; Score 273.4; DB 8; Length 3132;
Best Local Similarity 57.2%; Pred. No. 2.7e-53;
Matches 544; Conservative 11; Mismatches 372; Indels 24; Gaps 3;

QY 95 GGAGATGGTGAAGGGCAGCCGTTCCAGCTGGGCGCGCTACACGCAATGTCAGTGTGACGTACAT 154
DB 41 GCAGGAGCTGAACAAGACCGTCTGGGAGGTGCCCGAGCGATACCAGAACCTGTCCCGCGT 100

QY 155 CGCGAGGGCGGTACGGCATGCTACGCTGGCGCTATGACCAGCTGGCGCAAGACTCGCGT 214
DB 101 GGCGTCGGAGCGCTACGGCTCGGTGTGTGCTGTWYGAACACGCGGACATCGTGT 160

QY 215 GGCATCAAGAAGAT---CAGCCCTTCGAACATCAGACCTACTGCCAGCGCACGCTCG 271
DB 161 GGCAGTGAAGAGCTGTGAGACCGKTTAGYCCATCATTCAGGCCAAAGGWCCTACAG 220

QY 272 GGAGATCCAGATCCTGCTGGCTTCGCGCATGAGAAATGTCATCGGCATCCGAGACATTC 331
DB 221 GGAGCTCGCGCTGCTGAAGCACAATGAACACGAGAAATGTATGGTCTGTGTGATGTGT 280

QY 332 CGGGCGTCCA---CCCTGGAGCCATGAGAGTGTCTACATTTGTCAGGACTGTATGGA 388
DB 281 TACACCTGCAAGGTCCCTGGAGAAATTCAACGATGTGTACCTGGTGACCCATCTCATGG 340

QY 389 GACTGACCTGTACAACTTGTCTGAAGCCAGCAGCTGAGCAATGACATATCTGTACTT 448
DB 341 GGCAGACCTTGACAAACATCGTGAAGTGTGAGAGCTTACGATGACCAAGTTCAAGTTCT 400

QY 449 CTTCTACAGATCCTCGGGCGCTCAAGTACATCCACTCGGCAAGCTGCTCCACCGAGA 508
DB 401 TATCTACAGATCCTCGAGGGCTGAAGTATATACACTCGGCTGACATAATCCACAGGA 460

QY 509 TCTAAAGCCCTCAAACCTGCTCATCAACACCACCTGGACCTTGAAGTTTGTATTCGG 568
DB 461 CTTAAAGCCCGAACCTCGCTGAATGAAGACTGTGAGCTGAAGATTCTGTGATTTGG 520

QY 569 CTGGCCCGGATGTCGCGATCCTGAGCATGACCACACCGGCTTCTGACGGAGTATGTGC 628
DB 521 GCTGGCTCGGCACACTGAT-----GAGGAATGACCGGCTACGTGGC 562

QY 629 TACGCGCTGGTACCGGGCCCCAGAGATCATGTGAACTCCAAGGGCTATACCAAGTCCAT 688
DB 563 TACCCTGGTGTACAGAGCCCCGAGATTAAGTGAATGCTGAATGCTGATGCACTACCAACAGACAGT 622

QY 689 CGACATCTGGTCTGTGGCTGCAATTCGCTGAGATGCTCTTAAACGGCCCATCTTCC 748
DB 623 GGATATTTGGTCCGCTGGCTGCACTATGCTGAGCTGTGACCGGAAGAACGTTGTTCC 682

QY 749 TGGCAAGCACTACTCGATCAGCTCAACACATTTCTGGGCATCTGGGCTCCCGCATCCCA 808
DB 683 TGGTACAGACCATATTGATCAGTTGAGCTCATTTTAAGACTCGTTGGAAACCCCGGGGC 742

QY 809 GGAGGACCTGAATTGTATCATCAACATGAAGGCCCGGAAACTACTACATGCTCTGCGCCCTC 868
DB 743 TGAGCTTCTGAAGAAAATCTCTCTCAGAGTGTGCAAGAAACTACATTCAGTCTCTGGCCCA 802

QY 869 CAAGACCAAGTGGCTGGGCCAAGCTTTTCCCAAGTCAGCTCAGACTCCAAAGCCCTTCACCT 928
DB 803 GATGCCGAAGATGAATTCGCAAAATGTATTTATTGTGTGCAATTCCTCTGGCTGCSACCT 862

QY 929 GCTGGACCGGATGTTAACTCTTTAAACCCCAATAAACCAGATCACAGTGGAGGAGCGTGGC 988

DB 863 GCTGCAAAAAGATGCTGGTTTTTGGACTCRGATAAGAGGAWCACAGCAGCCCAAGCTCTTGC 922
QY 989 TCACCCTTACCTGGAGCAGTACTATGACCCGACGATGAGCCAGTGGCGGA 1039
DB 923 GCATGCTACTTTTGTCTCAGYACCAAGCCCTGATGATGAGCCAGTGGCTGA 973

RESULT 7

US-10-770-726-25
; Sequence 25, Application US/10770726
; Publication NO. US20050266409A1
; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Brown, Eugene

; APPLICANT: Liu, Wei

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

; FILE REFERENCE: AM101079 (031896-010000)

; CURRENT APPLICATION NUMBER: US/10/770,726

; CURRENT FILING DATE: 2004-02-04

; NUMBER OF SEQ ID NOS: 48640

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 25

; LENGTH: 1888

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-770-726-25

Query Match 14.9%; Score 273.2; DB 7; Length 1888;

Best Local Similarity 56.5%; Pred. No. 2.5e-53;

Matches 603; Conservative 0; Mismatches 438; Indels 27; Gaps 4;

QY 95 GGAGATGTGAAGGGCAGCCGTTGACGTGGGCGCGCTACACGCAATGTCAGTGTGACGTACAT 154
DB 131 GCAGGACGTCAACAAGACCGCTGGAGCTGCCAAGACCTACGTCTCCCCGACGACGT 190

QY 155 CGCGAGGGCGGTACGGCATGCTCAGCTTCGGCTTATGACCACTGACCACTGCGCGCAAGACTCGCGT 214
DB 191 CGCGAGGGCGGTATGGCTCGCTGCTCGGCCATCGACACGCTCAGGGGAGAGGT 250

QY 215 GGCATCAAGAAGATCAGC---CCCTTGAAACATCAGACCTACTGCGAGCGCACGCTCG 271
DB 251 GGCATCAAGAAGCTGAGCCGACCCCTTTCAGTCCGAGATCTTCGCCAAGCGCGCTACCG 310

QY 272 GGAGATCCAGATCCTGCTGGCTTCGCCATCAGATGTCATCGGCATCCGAGACCGAGAC 328
DB 311 GGAGCTGCTGCTGTGAAGACATGACATGAGAACGTCATTGGGCTCTGTGATGCTT 370

QY 329 TCTGGGGCGTCCACCTTGAAGCCATGAGAGATGTCTACATTGTGCGAGACCTGATGGA 388
DB 371 CACCCAGCGCTCTCCCTGGCAACTTCTATGACTTCTACCTGTGTGATGCCCTTCATGCA 430

QY 389 GACTGACCTGTACAAGTTGCTGAAAGCGAGCAGCTGAGCAATGACCATATCTGTACTT 448
DB 431 GACCGATCTCAGAAGATCATG---GGATGAGTTTCACTGAGGAGAGATCCAGTACCT 487

QY 449 CTTCTACAGATCCTGGGGCGCTCAAGTACATCCACTCGGCCAAGCTGCTCCACCGAGA 508
DB 488 GGTGATCAGATGCTCAAGGGCTTAAAGTACATCCACTCTGCTGGGGTCTGTGACAGGA 547

QY 509 TCTAAAGCCCTCCAACTGCTCATCAACACCACTGCGCACCTTAAGATTTGATTTCCG 568
DB 548 CTGGAAGCCAGCAACCTGGCTGTGATGAGGACTGTGAACCTGAAGATTTCTGATTTGG 607

QY 569 CTGGCCCGGATGCGGATCTCTGAGCATGACACCGGCTTCTGACCGAGTATGTGC 628
DB 608 GCTGGCGGACATGACAGCGCGAG-----ATGACTGGCTACGTGT 649

QY 629 TACGGCTGTTACCGGGCCCCCAGAGATCATGCTGAACCTCAAGGGCTATACCAAGTCCAT 688
DB 650 GACCCGCTGGTACCGAGCCCCCGAGGTGATCTCAGCTGGATGCACTACACACAGACT 709

QY 739 CCATCTTCCCTGCAAGCACTACTCGATCAGCTCAACCACTCTTCTGGGATCTCTGGCT 798
 Db 586 TCCTGTTCAAAGGCAATGACCACTGACCACTGCAAGGAGATCATGAAAGTCAAGGGA 645
 QY 799 CCCATCCAGAGGAGGACCTGAAATGTTATCATACATACATGAGGAGGAGGAGGAGG 858
 Db 646 CACCCCTCTGATGTTTGTACAGAGCTACAGAGTGTGAGGAGGAGGAGGAGGAGG 705
 QY 859 CTCTGCCCTCAAGACCAAGGCTGCTGGGCAAGCTTTTCCCAAGTCAAGTCCAAAG 918
 Db 706 GCTCTCTGAGTTGGAAGAGGATTTCTGCTGCTGCAATGCAAGCCCTCAGG 765
 QY 919 CCCTTGACCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 978
 Db 766 CCGTGAATCTCTGGAAGAGATGCTGCTGTTGATGCGGAAAGGAGGAGGAGGAGGAGG 825
 QY 979 AAGGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1038
 Db 826 AGGCAATTAAGCCCACTTCTGAGTCCCTTCCGAGACCTGAGGAGGAGGAGGAGGAGG 885
 QY 1039 AGGAGCCCTTCACTTCCGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1098
 Db 886 AGAATATGATGACTCTTGTATGACGTAGACCGCACCTTGTAGATGGAAGGCTGTTA 945
 QY 1099 TCATCTTCCAGGAGACAGCAGCTTCCAGCCGAGGAGGAGGAGGAGGAGGAGGAGG 1158
 Db 946 CGTATAAGGAAGTGTCTCAGCTTCAAGCCTCCAGGAGGAGGAGGAGGAGGAGGAGG 1005
 QY 1159 AGACATCTCTG 1169
 Db 1006 AGACAGCTCTG 1016

RESULT 11
 US-11-136-527-2977
 ; Sequence 2977, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2977
 ; LENGTH: 3824
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-11-136-527-2977

Query Match 11.1%; Score 203.8; DB 8; Length 3824;
 Best Local Similarity 56.5%; Pred. No. 2.7e-37;
 Matches 420; Conservative 0; Mismatches 317; Indels 6; Gaps 2;
 QY 336 GCGTCCACCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 395
 Db 672 GTGGGCTCTTAAACAGAGCTGAATAGCGTCTACATTTGTTACAGGAGTACATGGAGACAG 731
 QY 396 CTGTACAGTTGCTGAAAGCCAGCAGCTGAGCAATGACCATATCTGCTACTCTCTAC 455
 Db 732 TTGGGCAAGCTGTGGAGAGGGCCCTTTACTGGAGGAGCATGCGAGCTCTTCAATGTAC 791
 QY 456 CAGATCTCTGCGGGGCTCAAGTACATCCATCTCCGCAACGCTCTCCACCGAGATCTAAAG 515
 Db 792 CAGTGTCTGCTGGGCTCAAGTACATCCATCTCGAAGGCTGCTGCACAGGAGTCTCAAG 851
 QY 516 CCCTCCACCTGCTCATCAACAC---CACCTGCGACCTTAAGATTGTTGAGTTTCGGGCTG 572
 Db 852 CCGGCCAACCTTTTCATTAACACTGAGAGCTTGGTGTCTGAAGATTGGTGACTTTGGGCTG 911

QY 573 GCCCGGATTCGATCTCTGAGCATGACCAACCGGCTTCTGACGAGTATGTTGGCTAGC 632
 Db 912 GCCCGGATCATGGATCTCTCATTTATTTCCCATAGGGTCACTTTCTGAGGATTTGGTTACC 971
 QY 633 CGCTGTATCCGGGCCCCAGAGATCATGCTGAACTCCAAGGGCTATACCAAGTTCATTCGAC 692
 Db 972 AAATGGTACAGATCTCCAGCGCTTTTACTTTCTCTAATACTATATCTAAGCCATTGAC 1031
 QY 693 ATCTGTCTGTGGGCTGCAATTCGTGTGAGATGCTCTTAAACCGGCCCATCTTCCCTGGC 752
 Db 1032 ATGTGGGCTGAGGCTGCACTTTGCTGAAATGCTGACTGGTAAACCCCTCTTTGAGGT 1091
 QY 753 AAGCACTACTGATCAGCTCAACCACTTCTGGGCACTCTGGGCTCCCATCCAGGAG 812
 Db 1092 GCACATGAATTTGAACAGATGCACTGATCTTTGGAGTCTATCCCTGTTGTGACGAGGAA 1151
 QY 813 GACCTGAATTTGATCATCAACATGAAGGCCGAAACTACCTACAGTCTCTGCCCTCCAAAG 872
 Db 1152 GATCGGAGGAGGCTCTCAGCGATGTTCCAGTTTACATTAGAAACGACATG---ACTGAG 1208
 QY 873 ACCAAGGTGGCTTGGGCAAGCTTTTCCCAAGTCAAGTCCAAAGCCTTGACCTGTG 932
 Db 1209 CCACACAAACCGCTGACTCAGCTGCTTCGGGGAATTAGTCCGGAGCACTGGATTTCTCTG 1268
 QY 933 GACCGATGTTTAACTTTAAACCCCAATAACCGATCAAGTGGAGGAGGAGGAGGAGGAGG 992
 Db 1269 GAACAGATTTCTGACGTTTCACTGCTGACCGCTGACAGCCGAGGAGGAGGAGGAGGAGG 1328
 QY 993 CCTTACTCTGGAGCACTATGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1052
 Db 1329 CTTTACATGAGCATCTACTCTTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1388
 QY 1053 TTGCGCATGAGCTGGATGACCT 1075
 Db 1389 ATAGAAGACGAGTGGACGACAT 1411

RESULT 12
 US-11-136-527-3508
 ; Sequence 3508, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 3508
 ; LENGTH: 822
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-11-136-527-3508

Query Match 9.1%; Score 166.4; DB 8; Length 822;
 Best Local Similarity 56.8%; Pred. No. 6.2e-29;
 Matches 349; Conservative 0; Mismatches 256; Indels 9; Gaps 2;
 QY 355 TGAGAGATGTCTACATTTGTCAGGAGCCTGATGGAGAGTCACTGTGTACAGTTGCTGAAAA 414
 Db 215 TCAGCGTGGCTTATTTGTTCCAGAGTACATGAGAGTACCTGGCATGCTGTCTAGAGC 274
 QY 415 GCCAGAGCTGAGCAATGACCATATCTGTACTTCTTACAGATCTCTGGGGGCTCA 474
 Db 275 AGGCGAGCTGACCGGAGGAGCAGCTAAAGCTATTTCATGTACACAGCTGCTGGGCTCA 334
 QY 475 AGTACATCCACTTCGCGCAAGCTGCTCCAGGAGATCTAAAGCCCTCCAACTGCTCATCA 534

Db 335 AGTACATCACTCTGCAACGCTGTGACACAGGACCTGAAGCCCGCAACATTTTCATCA 394
QY 535 ACAC---CACCTGGACCTTAAGATTGTGATTTGGCCCTGGCCCGGATTCGCGATCCTG 591
Db 395 GCACGGAGGACCTCGTCTGAAGATCGGGGATTTGCGGCTGGCCAGAAATCGCGGACAGC 454
QY 592 AGCATGACACACCGGCTTCTGACGAGAGTATGTGGCTACGCGCTGATACCGGGCCCCAG 651
Db 455 ATTACTCCCAAGGGGTATCTGTGGAAGGGTGTGTGCAAAAGTGGTACCGCTCTCCAC 514
QY 652 AGATCATGTGAACTCCAAAGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCA 711
Db 515 GACTGTCTGTGCCGCAACTACACGAAGCCATCGATATGTGGCAGCTGGCTGCA 574
QY 712 TCTGGCTGAGTCTCTTAACCGGCCCACTTCCTCGTGGCAAGCACTACCTGGATCAGC 771
Db 575 TCCTAGCGGAGATGCTCACGGGAAATGCTCTTTGCTGGGGCTCACGAGCTTGAGCAGA 634
QY 772 TCAACACATTTCTGGGCATCTGGGCTCCCATCCAGGAGGACCTGAATTGTATCATCA 831
Db 635 TGCAGCTCATCTTAGACACCATCCCTGTAGTGGGAGGAGAACAGGAGGAGCTGCTCA 694
QY 832 ACATGAAGGCCGAAACTACTACAGTCTCTGCCCCTCCAAGACCAAGGTGGCTTGGGCCA 891
Db 695 GGGTGTGTC---GTCTTTGTGACGAGCACTTTGGGAGGTGAAGAGGCCACTGCGTA 748
QY 892 AGCTTTTCCCAAGTCAGACTCAAAGCCCTTGACCTGTGTCGACCGAGTGAACCTTTA 951
Db 749 AGCTACTCCCGGATGTCAACCGTGAAGCCATTGACTTCTTGGAGAAGATCCTGACGTTCA 808
QY 952 ACCCCCAATAACGG 965
Db 809 GCCCATGGACCGG 822

RESULT 13

US-11-136-527-238
; Sequence 238, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 238
; LENGTH: 1980
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-238

Query Match 9.0%; Score 165.8; DB 8; Length 1980;
Best Local Similarity 55.7%; Pred. No. 1.1e-28;
Matches 452; Conservative 0; Mismatches 332; Indels 27; Gaps 6;

QY 115 CGTTCGACGTGGGCGCGCTACACGAGTTGCAGTATCATCTCGCGAGGGCGGTACGGCA 174
Db 419 CTTTCAGGTTCTAAGCGCTACAGAACCTGAAGCGATCGGCTCTGGGCTCAGGGAA 478
QY 175 TGGTCACTCGGCTATGACCAAGTGCAGCAAGACTCGCTGGCCATCAAGAAGATCAGC- 233
Db 479 TAGTTTGTGCTGATGACGCTGTCTCGACAGMAATGTGGCCATTAAGAAGCTCAGCA 538
QY 234 --CCCTTCGAACATCAGACCTACTGCGCAGCGCAGCTCCCGGAGATCCAGATCCTGCTGC 291
Db 539 GACCTTTCCAGAACCAAACTCATGCCAAGAGGGCTTACCGGGAGCTGGTCTCATGAAGT 598
QY 292 GCTTCGCCCATGAGAATGTATCGGCATCCGA---GACATTTCTCGCGGCGCTCCACCGCTGG 348

Db 599 GTGTGAACCAATAAACAATTATTAGCTTATAAATGTCTTTTACACCCAGAAAAACACTGG 658
QY 349 AAGCCATGAGAGATGTCTACATTGTGAGGACCTGATGGAGACTGACCTGTACAAAGTTGC 408
Db 659 AGGAGTTCCAAGATGTTTACTTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 718
QY 409 TGAAGAAGCCAGCAGCTGAGCAATGACCATATCTGCTACTTCTCTACAGATCTCTGCGG 468
Db 719 TTCAAG---TGAGCTGGACCAAGCGGATGTCGTACTTGTGTACAGATGCTGTGCGG 775
QY 469 GCCTCAAGTACATCACTCCGCAACGCTCTCCAGGAGATCTAAAGCCCTCCAACTGTC 528
Db 776 CGATCAACACCTCCACTCCGCTGGGATCATCCAAGGAGCTTAAAAACCCAGTAACATCG 835
QY 529 TCATCAACACCACTGCGACCTTAAGATTGTGATTTGGCTCTGCGCCCGGATGTCGGATC 588
Db 836 TAGTCAAGTCTGATTGCACTGAAATCTCTGGACTTTGGACTGGCCAGGACAGC----- 890
QY 589 CTGAGCATGACACACCGGCTTCTGAGCGGAGTATGTGGCTACGCGCTGGTACCGGGCC 648
Db 891 -----GGGCAAGCTTCATGATGATCCCGTATGTGTGTGACGAGATATTACAGAGCC 943
QY 649 CAGAGATCATGCTGAACCTCAAAGGCTATPACCAAGTCCATCGACATCTGCTGTGGGCT 708
Db 944 CCGAGTCACTCG---GGCATGGGCTACAAGGAGAACTGGGACATATGCTCTGTGGGCT 1000
QY 709 GATTTCTGGCTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATC 768
Db 1001 GCATCATGGGAGAAATGGTTCGTCAAAATCTCTTCCCGGAAAGGAGTATATTGACC 1060
QY 769 AGCTCAACACATTTCTGGGCTCTGGGCTCCCCATCCAGGAGGACCTGAATTGTATCA 828
Db 1061 AGTGAACCAAGTCAAGAGGAGCTAGGAACTCCGTGTCCAGATTCATGAA---GAAAT 1117
QY 829 TCAACATGAAGGCCCGCAACTACTACAGTCTCTGCCCTCCCAAGCAAGGTGGCTTGG 888
Db 1118 TGCAGCCCACTGACAGAACTAGTGGAGAACCGGCCCAAGTATGAGCGCCTCACCTTCC 1177
QY 889 CCAAGCTTTTCCCAAGTCAAGCTCAAAAGC 919
Db 1178 CCAAGCTTTTCCAGATTCCTCTTCCCGCAGC 1208

RESULT 14

US-11-136-527-7604
; Sequence 7604, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7604
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7604

Query Match 9.0%; Score 165; DB 8; Length 600;
Best Local Similarity 57.0%; Pred. No. 1.2e-28;
Matches 344; Conservative 0; Mismatches 250; Indels 9; Gaps 2;

QY 366 TACATTGTGCAGGACCTGATGGAGACTGACCTGTACAAAGTTGCTGAAAGCCAGCAGCTG 425
Db 4 TATATTGTCCAAGAGTATCATGAGAGTGAACCTGGCATGCTGCTAGAGCAGGACGCTG 63

```
QY 426 AGCAATGACCATATCTGTCTACTTCTCTACAGATCTCGGGGCTCAAGTACATCCAC 485
Db 64 ACCGAGGAGCAGCGTAAGCTATTTCATGTACAGCTGCTGGGTGCTCAAGTACATCCAC 123
QY 486 TCCGCAACCTGCTCCACAGAGCTTAAAGCCCTCCAACTGCTCATCAACAC--CACC 542
Db 124 TCTGCCAACGTGTTCACAGGACCTGAAGCCCGCCAACTTTTCATCAGACCGAGGAC 183
QY 543 TGGACCTTAAAGATTGTGATTTTCGGCTCGCCCGGATTCGCCGATCCTGAGCATGACCAC 602
Db 184 CTGCTGCTGAAGATCGGGGATTTTCGGCTCGCCAGAACTCGGACCACTTACTCCAC 243
QY 603 ACCGGCTTCCTGACGAGATATGTGGCTACCGGCTGTGTACCGGGCCCGAGATCATGCTG 662
Db 244 AAGGGTTATCTGTCCGAAGGGTTGGTGACAAAGTGGGTACCGCTCTCCACGACTGCTCCTG 303
QY 663 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTGAG 722
Db 304 TCCCCGAACAACTACACGAAGCCATCGATATGTGGGACCTGGCTGCATCTTAGCGGAG 363
QY 723 ATGCTCTCTTAACGGGCCATCTTCCCTGGCAAGCACTTACCCTGGATCAGCTCAACCACTT 782
Db 364 ATGCTACGGGGAAATGCTCTTTGTCTGGGCTCACGAGCTTGAGCAGATGCAGCTCATC 423
QY 783 CTGGGCATCTGGGCTCCCATCCAGAGGAGCTGAATTTGTATCATCAACATGAAGGCC 842
Db 424 CTAGACACCATCCTGTGTAGTGGGAGGAGAAAGAGGAGCTGTCTCAGGGTGTATGCC- 482
QY 843 CGAACTACCTACATGCTCTGCCCTCCAGACCAAGTGGCTTGGGCCAGCTTTTCCCC 902
Db 483 -----GTCTTTGTACAGCACTTGGGAGGTGAAGAGGCCACTGCGTAAGCTACTCCCG 537
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGCTCGACCGGATGTTAACTTTAAACCCCAATAA 962
Db 538 GATGTCAACCGTGAGGCCATTGACTTCTGAGAGAGATCCTGAGGTTAGCCCCATGGAC 597
QY 963 CGG 965
Db 598 CGG 600
```

RESULT 15

```
US-11-136-527-4380
; Sequence 4380, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4380
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (587)..(587)
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-4380
```

```
Query Match 8.9%; Score 164.2; DB 8; Length 600;
Best Local Similarity 65.2%; Pred. No. 1.8e-28;
Matches 362; Conservative 6; Mismatches 164; Indels 23; Gaps 8;

QY 1284 ACTGTCCGAGCCCGGACCTTGGCAGCCAGCCGGGTGGAGCATCGGGCTGGCCACCT 1343
Db 10 ACTTGTCTCAACCGGACCCCGGAGCCAGG-CTGGACCAAGGGTGGGCTGGCACCCC 68
```

```
QY 1344 CTCCTCTTTGCTGAGGCTTCCAGCTTTCAGGAGGCCAAGGCTTCTCTCCCAACCCGCC 1403
Db 69 TCTCACTCTGCTGGGTCTCTCTSSITTCAGAGG--CTTCTCCCACTCCAGTCCCTGCC 126
QY 1404 CTCCCCACGGGCTCGGGAGCTCAGGTGGCCCCAGTTCAAATCTCCCGTGTCTGTCTG 1463
Db 127 CATCTCCCCTTGACCTGATGATGAGTGGTCCAGAGCTGATCTCTGTCTGTCTCTT 186
QY 1464 CGCCCTTACCTTTCCCAGCGTCCCAGTCTCTGGCAGTTCTGGGAATGGAAGGTTCTGGCT 1523
Db 187 TATCTATCTCTGTAGCCCCAGCTCTGGTA-GACGGTTCTGGAAATGGAAGGCTATGACC 245
QY 1524 GCCCCAACTGCTGAAGGGCAGAGGTGGAGGTGGGGGGGCTGAGTAGGGACTCAGGGC 1583
Db 246 GGCCTA-----GGACCTGTGCTACAGAGGGGTGGAGGGCACTGAGT-----AGGCT 291
QY 1584 CATGCTGCCCCCTCTCATCTCAATTCAAACCCCAACCTAGTTTCCCTGAAAGGAACATTCTT 1643
Db 292 AAGCTCTGCCCTACTCATCTGTGGAAACCCCAACCCCAATTTTCCCTGACAGAACATTCTT 351
QY 1644 TAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCGGCCGAATCCCTTCCCTGTCAAA 1703
Db 352 AAATCTCAAGGGCTAGTTTCCCTGAGGAGCCAGCTAGGCTTAAACCTCTCTCTCTC-AA 410
QY 1704 GCTGTCACTTCCGGTGCCTCGCTGCTTCTGTGTGT-GGTGAAGCAGAGTGGAGCTGGGG 1762
Db 411 GCTGCCACATGTAAACGCCCKTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470
QY 1763 GGGGTGGAGCCCGGCGCCCTGCCACCTCCCTGACCC--GTCTAATATATAAATATA 1819
Db 471 CCGTGGAGAGCCCGGCGCCCTGCCACCTCCCTGCTGTATCTAATATATAAATATA 530
QY 1820 GAGATGTCTATGG 1834
Db 531 GAGATGTCTATATGG 545
```

Search completed: February 7, 2006, 14:11:06
Job time : 796.931 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:19:31 ; Search time 266.938 Seconds
(without alignments)
551.408 Million cell updates/sec

US-10-623-108-2

Title:
Perfect score: 1763

Sequence: 1 MAAAAAQQGGGGPRRTG.....LKELIFQTARFQGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq 21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	335	8	ADH48364 Human KPP
2	1763	100.0	335	8	ABM82550 Human dia
3	1763	100.0	335	8	ABM82548 Human dia
4	1763	100.0	335	9	ADW12901 Human SMA
5	1731	98.2	379	2	AAW15506 Mitogen a
6	1731	98.2	379	4	AAM38754 Human pol
7	1731	98.2	379	6	ABR43898 Human ERK
8	1731	98.2	379	7	ADF45050 Human kin
9	1731	98.2	379	7	ADH59632 Erk1 prot
10	1731	98.2	379	8	ADO22522 Biochemic
11	1731	98.2	379	8	ADQ37846 Human pho
12	1731	98.2	379	9	ADV42070 Human ext
13	1731	98.2	379	9	ADW12905 Human SMA
14	1731	98.2	379	9	ADZ65036 Mitogen a
15	1731	98.2	403	4	AAM40540 Human pol
16	1725	97.8	379	2	AAW42413 Extracell
17	1725	97.8	379	4	AAW67618 Amino aci
18	1725	97.8	379	4	AAW67439 Human ext
19	1725	97.8	379	5	ABG31847 Human ext
20	1725	97.8	379	6	ABU89742 Protein d
21	1725	97.8	379	7	ABR62625 Human pro
22	1725	97.8	379	8	ADO24425 Human PRO
23	1725	97.8	379	9	ADZ10062 Human bre
24	1725	97.8	631	2	AAW85006 Erk1-gree

RESULT 1
ADH48364
ID ADH48364 standard; protein; 335 AA.
XX ADH48364;
AC ADH48364;
XX XX
DT 25-MAR-2004 (first entry)
XX XX
DE Human KPP protein SEQ ID NO:22.
XX XX
KW human; KPP; kinase; phosphatase; enzyme; cytostatic; anorectic;
KW immunosuppressive; KPP-Antagonist; KPP-Agonist; gene therapy;
KW autoimmune disorder; obesity; cancer.
XX OS Homo sapiens.
XX XX
PN WO2004001008-A2.
XX XX
PD 31-DEC-2003.
XX XX
PF 19-JUN-2003; 2003WO-US019660.
XX XX
PR 21-JUN-2002; 2002US-0390652P.
PR 15-JUL-2002; 2002US-0396196P.
(INCY-) INCYTE CORP.

Marquis JP, Baughn MR, Tran UK, Hafalia AJA, Kable AE, Emerling BM, Elliott VS, Lindquist EA, Richardson TW, Khare R, Swarnakar A, Lee SY, Ramkumar J, Chawla NK, Becha SD, Mason PM, Hawkins PR, Bulloch SA, Jin P, Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
WPI; 2004-082489/08.
N-PSDB; ADH48413.
New human kinases and phosphatases (KPP) polypeptide, useful for preparing a composition for treating a disease associated with decreased expression or overexpression of functional KPP e.g., cancer.
Claim 1; SEQ ID NO 22; 336pp; English.
The invention relates to novel isolated human kinases and phosphatases (KPP) polypeptides. A protein of the invention has cytostatic, anorectic, and immunosuppressive activity, and acts as a KPP-Antagonist, or KPP-Agonist. A polynucleotide of the invention may have a use in gene therapy. The polypeptide is useful for preparing a composition for


```

PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthehorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kifton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patry S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR N-PSDB; ACN41200.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 335 AA;
Query Match 100.0%; Score 1763; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.2e-174;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAAAGGGGGPRTEGVGPGVGEVEMVKGQPFDPGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRTEGVGPGVGEVEMVKGQPFDPGPRYTQLOVIGEGAYGMVSSAY 60
QY 61 DHVKTRVAIKKISPFHEQTYCQRTLEIRIQILFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPFHEQTYCQRTLEIRIQILFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGUKYTHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGUKYTHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
DB 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
QY 301 PFTFAMELDDLPKRLKELIFQETARFQGVLEAP 335
DB 301 PFTFAMELDDLPKRLKELIFQETARFQGVLEAP 335

```

```

RESULT 4
ADW12901
ID ADW12901 standard; protein; 335 AA.
XX
AC ADW12901;
XX
DT 07-APR-2005 (first entry)
DE Human SNAPK3V1 variant protein.
XX
KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW enzyme.
XX
OS Homo sapiens.
XX
PN US2005013817-A1.
XX
PD 20-JAN-2005.
XX
PF 18-JUL-2003; 2003US-00623108.
XX
PR 18-JUL-2003; 2003US-00623108.
XX
PA (DAIK/) DAI K.
XX
PI Dai K;
XX
DR WPI; 2005-080923/09.
DR N-PSDB; ADW12900.
XX
XX New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
PT diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
PT in a mammal.
XX
PS Claim 1; SEQ ID NO 2; 55pp; English.
XX
XX The invention relates to novel isolated polypeptide (I) comprising an
CC amino acid sequence selected from sequences comprising 335 or 359 amino
CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
CC associated with the deficiency of the SNAPK3 gene in a mammal. The
CC fragments of the polypeptides and polynucleotides can also be used as
CC primers or probes. This sequence corresponds to the SNAPK3V1 variant
CC protein.
XX
SQ Sequence 335 AA;
Query Match 100.0%; Score 1763; DB 9; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.2e-174;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAAAGGGGGPRTEGVGPGVGEVEMVKGQPFDPGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRTEGVGPGVGEVEMVKGQPFDPGPRYTQLOVIGEGAYGMVSSAY 60
QY 61 DHVKTRVAIKKISPFHEQTYCQRTLEIRIQILFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPFHEQTYCQRTLEIRIQILFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGUKYTHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGUKYTHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
DB 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
QY 301 PFTFAMELDDLPKRLKELIFQETARFQGVLEAP 335

```

Db 301 PTFANELDDLPKRLKELIFQETARFQGVLEAP 335

RESULT 5
AAW15506
ID AAW15506 standard; peptide; 379 AA.
XX AAW15506;
DT 03-JUN-1997 (first entry)
XX Mitogen activating protein kinase ERK1.
XX Antibody; MAP kinase; human; mitogen activated protein, lymphoid cell;
KW recognition sequence; IgG; immunogen; rat; ERK2; hybridoma; spleen cell;
KW MAP kinase-related disease; cancer-associated disease; brain tumour;
KW metabolic disorder; diabetes mellitus; circulatory disease; rheumatism;
KW arteriosclerosis; allergic diseases; central nervous system disease;
KW asthma; Alzheimer's disease; Parkinson's disease; senile dementia; ERK1;
KW bone/joint disease; pollenosis; atopic dermatitis.
XX Homo sapiens.
XX EP735370-A1.
XX 02-OCT-1996.
XX 28-MAR-1996; 96EP-00104966.
XX 28-MAR-1995; 95JP-00070125.
XX 24-NOV-1995; 95JP-00305456.
XX 08-DEC-1995; 95JP-00320577.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Tani A, Ichimori Y;
XX WPI; 1996-444623/45.
XX New antibodies to human MAP kinase - used for the detection, assay and
PT purification of various MAP kinase species.
XX Example 3; Page 34-35; 59pp; English.
XX This sequence represents the human mitogen activated protein (MAP) kinase
CC ERK1. Fragments of this sequence (such as AAW15501) are recognition
CC sequences for the antibodies of the invention. The antibodies of the
CC invention are IgG type monoclonal antibodies (Mab) generated by using
CC human ERK1 as an immunogen and are capable of binding to ERK1 at the
CC sites represented by these peptides. However, the Mab is incapable of
CC binding to human or rat MAP kinase ERK2, or rat MAP kinase ERK1. The Mab
CC are produced from cloned hybridomas derived from spleen cells of a mammal
CC immunised with human ERK1 and homologous or heterologous lymphoid cells.
CC The antibodies can be used for detecting or assaying MAP kinases. They
CC can be used for diagnosing MAP kinase-related diseases, such as cancer-
CC associated diseases (e.g. brain tumour), metabolic disorders (e.g.
CC diabetes mellitus), circulatory disease (e.g. arteriosclerosis), allergic
CC diseases (e.g. asthma, pollenosis, atopic dermatitis), central nervous
CC system diseases (e.g. Alzheimer's disease, Parkinsonism, senile dementia)
CC and bone/joint diseases (e.g. rheumatism). They can also be used to
CC investigate the role of MAP kinases in-vivo and the mechanism of drug
CC action. The antibodies can also be used for purifying MAP kinases. By
CC using antibodies specific for various types of MAP kinase, the activity
CC of any given species of MAP kinase can be detected or assayed with high
CC sensitivity to the exclusion of other species of MAP kinase
XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 2; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRRTTEGVGPGVEVMWKGQPFVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGEPRRTTEGVGPGVEVMWKGQPFVGPRTYQIQYIGEGAYGMVSSAY 60
QY 61 DHVTRTRVAIKKISPFHQTYCQRTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
DB 61 DHVTRTRVAIKKISPFHQTYCQRTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYOILRGLKXIHSANVLRDLKPSNLLINTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYOILRGLKXIHSANVLRDLKPSNLLINTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRYRAPEIMLNSKGYTKSIDISWVCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRYRAPEIMLNSKGYTKSIDISWVCILAEMLS 240
QY 241 NRPIFGKHYLDQLNHIL----- 258
DB 241 NRPIFGKHYLDQLNHIL----- 258
QY 259 --ALDLLDRMLTFNPKRITVEEALAHPLYEQYDPTDEPVAEEPTFAMELDDLPKREL 316
DB 301 SKALDLLDRMLTFNPKRITVEEALAHPLYEQYDPTDEPVAEEPTFAMELDDLPKREL 360
QY 317 KELIFQETARFQGVLEAP 335
DB 361 KELIFQETARFQGVLEAP 379
RESULT 6
AAW38754
ID AAW38754 standard; protein; 379 AA.
XX AAW38754;
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 1899.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX N-PSDB; AAI57910.

PT Novel nucleic acids and polypeptides, useful for treating disorders such
 XX as central nervous system injuries.

XX Example 3; SEQ ID NO 1899; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 4; Length 379;
 Best Local Similarity 88.4%; Pred. No. 1.3e-170;
 Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
 DB 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
 QY 61 DHVRKTRVAIKKISPFPHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
 DB 61 DHVRKTRVAIKKISPFPHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
 QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKXTHSANVLRDLKPSNLLINTTCDL 180
 DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKXTHSANVLRDLKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
 DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
 QY 241 NRPIFGKHVLDQLNHIL----- 258
 DB 241 NRPIFGKHVLDQLNHILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPKSD 300
 QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 316
 DB 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 360
 QY 317 KELIFOETARFQPGVLEAP 335
 DB 361 KELIFOETARFQPGVLEAP 379

RESULT 7
 ABR43898

ID ABR43898 standard; protein; 379 AA.

XX AC ABR43898;

DT 11-AUG-2003 (first entry)

XX Human ERK1 protein.

XX DARPP-32; psychotic disorder; intracellular signaling protein;
 KW dopamine- and cAMP-regulated phosphoprotein; ERK1; ERK2; CREB;
 KW extracellular signal-regulated protein kinase; phosphorylation;
 KW cAMP-response element binding protein; schizophrenia; human.

XX Homo sapiens.

XX

PN WO2003021225-A2.

XX 13-MAR-2003.

XX 03-SEP-2002; 2002WO-US027802.

XX 31-AUG-2001; 2001US-0316338P.

XX (UYRQ) UNIV ROCKEFELLER.

XX (KARO-) KAROLINSKA INST.

XX Greengard P, Pisone G;

XX WPI; 2003-300912/29.

XX Identifying agent to be tested for ability to treat psychotic disorder,
 PT by contacting cells/tissues with candidate drug, determining
 PT phosphorylation levels of intracellular signaling proteins DARPP-32,
 PT ERK1, ERK2, CREB.

XX Disclosure; Page 77-78; 79pp; English.

XX The invention relates to identifying an agent to be tested for ability to
 CC treat psychotic disorder in patient. The method involves contacting cells
 CC or tissues with a candidate drug, and determining levels of
 CC phosphorylation of intracellular signaling proteins DARPP-32 (dopamine-
 CC and cAMP-regulated phosphoprotein), ERK1 and ERK2 (extracellular signal-
 CC regulated protein kinases 1 and 2), and CREB (cAMP-response element
 CC binding protein). The method is useful for identifying an agent to be
 CC tested for an ability to treat a psychotic disorder such as schizophrenia
 CC in a patient in need of such treatment. The present sequence represents a
 CC human ERK1 protein

XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 6; Length 379;
 Best Local Similarity 88.4%; Pred. No. 1.3e-170;
 Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
 DB 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
 QY 61 DHVRKTRVAIKKISPFPHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
 DB 61 DHVRKTRVAIKKISPFPHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
 QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKXTHSANVLRDLKPSNLLINTTCDL 180
 DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKXTHSANVLRDLKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
 DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
 QY 241 NRPIFGKHVLDQLNHIL----- 258
 DB 241 NRPIFGKHVLDQLNHILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPKSD 300
 QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 316
 DB 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 360
 QY 317 KELIFOETARFQPGVLEAP 335
 DB 361 KELIFOETARFQPGVLEAP 379
 RESULT 8
 ADF45050
 ID ADF45050 standard; protein; 379 AA.
 XX AC ADF45050;

XX 12-FEB-2004 (first entry)
XX Human kinase ERK1.
XX Human; protein kinase; enzyme; inhibitor; ERK1.
XX Homo sapiens.
XX WO2003081210-A2.
XX 02-OCT-2003.
XX 20-MAR-2003; 2003WO-US008725.
XX 21-MAR-2002; 2002US-0366892P.
XX (SUNE-) SUNESIS PHARM INC.
XX Prescott JC, Braisted A;
XX WPI; 2003-865136/80.
XX Identifying ligand binding to inactive conformation of target protein kinase (T) comprises contacting the conformation modified (T) which contains reactive group at binding site, with ligands and detecting kinase-ligand conjugate formation.
XX Disclosure; SEQ ID NO 19; 260pp; English.
XX The present invention relates to a method for identifying a ligand (L), which binds to an inactive conformation of target protein kinase (T). The method involves contacting inactive conformation of (T), which contains or is modified to contain a reactive group at or near a binding site of interest, with one or more ligand candidates capable of covalently bonding to the reactive group thus forming a kinase-(L) conjugate (C). The method is useful for identifying protein kinase inhibitors that preferentially bind to inactive conformation of a target protein kinase. The present sequence is a protein kinase which may be modified via an amino acid substitution, for use in the method of the invention.
XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 7; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAQQGGGGPRRTGEGVPGVGEVEMVKGQPDFVGPRTYQLOYGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGPRRTGEGVPGVGEVEMVKGQPDFVGPRTYQLOYGEGAYGMVSSAY 60
QY 61 DHVTRKTRVAIKKISPEFHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVTRKTRVAIKKISPEFHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLRIADPEHDHTGFLTYVATRWYRAPEIMLNSKGYTKSIDIVSWGICLAEMLS 240
DB 181 KICDFGLRIADPEHDHTGFLTYVATRWYRAPEIMLNSKGYTKSIDIVSWGICLAEMLS 240
QY 241 NRPIFFGKHLYDQLNHIL----- 258
DB 241 NRPIFFGKHLYDQLNHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVNAKLFPKSD 300
QY 259 --ALDILDLMTNPNKRIITVEALAHPLYEQYDPTDPEVAEPTFEAMELDDLPERL 316
DB 301 SKALDLDLMTNPNKRIITVEALAHPLYEQYDPTDPEVAEPTFEAMELDDLPERL 360
QY 317 KELIFOETARFQGVLEAP 335
|||||

DB 361 KELIFOETARFQGVLEAP 379
RESULT 9
ADH59632
ID ADH59632 standard; protein; 379 AA.
XX ADH59632;
XX ADH59632;
DT 25-MAR-2004 (first entry)
XX Erk1 protein.
XX cardiac disease; Raf-1; MEK1; Cardiant; telethonin; &agr;-actinin;
KW hypertrophy; MEK2; Erk1; Erk2.
XX Homo sapiens.
OS WO2003025205-A2.
PN 27-MAR-2003.
PD 18-SEP-2002; 2002WO-EP010489.
PF 19-SEP-2001; 2001US-0323566P.
PR 24-SEP-2001; 2001US-0324625P.
XX (MEDI-) MEDIGENE AG.
PA Nave B, Roenicke V, Leclair S, Funk M, Reuner B, Brinkmann K;
PI Henkel T;
XX WPI; 2003-371821/35.
DR N-PSDB; ADH59631.
XX Identifying and/or obtaining a compound useful for preventing or treating cardiac diseases, particularly congestive heart failure, comprises quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending on the compound.
XX Claim 33; SEQ ID NO 8; 41pp; English.
XX The present invention relates to identifying and obtaining a compound useful in the prevention or treatment of cardiac diseases, comprising quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending on the compound. The method is useful for identifying and/or obtaining compounds that may be used in the prevention or treatment of cardiac diseases, particularly congestive heart failure. The compound or protein is also used for the preparation of a pharmaceutical composition for prevention or treatment of a disease related to hypertrophy or impaired or increased activation of telethonin (T-Cap), &agr;-actinin, MHC, actin, titin, myomesin, nebulin, tropomyosin, troponin, Erk1/2 and/or MLCK. The protein or antibody which specifically recognizes the activated/phosphorylated form of the above polypeptide, is used for the preparation of a composition for diagnosing a disease or a predisposition for a disease related to hypertrophy or related to impaired or increased activation of Raf-1, MEK1/2 and/or Erk1/2. The present sequence represents Erk1.
XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 7; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAQQGGGGPRRTGEGVPGVGEVEMVKGQPDFVGPRTYQLOYGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGPRRTGEGVPGVGEVEMVKGQPDFVGPRTYQLOYGEGAYGMVSSAY 60
QY 61 DHVTRKTRVAIKKISPEFHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVTRKTRVAIKKISPEFHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

Qy 121 VQDLMTDLYKLLKSQLNDHICFYQLRGLKYYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQLNDHICFYQLRGLKYYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPIMLSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPIMLSKGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHIL----- 258
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPKSD 300
Qy 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 316
Db 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 360
Qy 317 KELIFOETARFQPGVLEAP 335
Db 361 KELIFOETARFQPGVLEAP 379

RESULT 10
ID ADO22522 standard; protein; 379 AA.
XX AC ADO22522;
DT 12-AUG-2004 (first entry)
DE Biochemical pathway-related human ERK protein SeqID8.

XX biochemical pathway; mammalian glioma tumour;
KW epidermal growth factor receptor; EGFR; mTOR; polypeptide inhibitor;
KW cancer; ERK; human.
XX Homo sapiens.

XX WO2004044218-A2.
XX 27-MAY-2004.
XX 05-NOV-2003; 2003WO-US035115.
XX 05-NOV-2002; 2002US-0423777P.
XX (REGC) UNIV CALIFORNIA.
XX (CELL-) CELL SIGNALING TECHNOLOGY INC.

XX Mischel PS, Sawyers CL, Smith BL, Crosby K;
XX WPI; 2004-411736/38.
XX Use of biochemical pathways associated with glioblastoma for, e.g.
PT identifying a mammalian glioma tumor that is likely to respond to an
PT epidermal growth factor receptor polypeptide inhibitor or an mTOR
PT polypeptide inhibitor.
XX Claim 1; SEQ ID NO 8; 66pp; English.

XX This invention relates to a novel use of biochemical pathways for
CC identifying a mammalian glioma tumor that is likely to respond to an
CC epidermal growth factor receptor (EGFR) polypeptide inhibitor or an mTOR
CC polypeptide inhibitor, or identifying a mammalian glioma tumor that does
CC not express or expresses a PTEN polypeptide and which is likely to
CC respond or not likely to respond to an inhibitor of mTOR polypeptide
CC activity or inhibitor of EGFR polypeptide activity, respectively. The
CC biochemical pathways are, in particular, disregulated in pathologies such
CC as cancer. The present sequence is that of the human ERK protein which
XX may be used in the method of the invention.

XX Sequence 379 AA;
XX Query Match 98.2%; Score 1731; DB 8; Length 379;

Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
Qy 1 MAAAAAOGGGGGPRTEGVGPGEVENVKQPPDVGPRYTQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAOGGGGGPRTEGVGPGEVENVKQPPDVGPRYTQLOYIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHOTYQRTLRQIILRFHENVIGIRDIRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHOTYQRTLRQIILRFHENVIGIRDIRASTLEAMRDVYI 120
Qy 121 VQDLMTDLYKLLKSQLNDHICFYQLRGLKYYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQLNDHICFYQLRGLKYYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPIMLSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPIMLSKGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHIL----- 258
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPKSD 300
Qy 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 316
Db 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 360
Qy 317 KELIFOETARFQPGVLEAP 335
Db 361 KELIFOETARFQPGVLEAP 379
RESULT 11
ADQ37846
ID ADQ37846 standard; protein; 379 AA.
XX AC ADQ37846;
XX 09-SEP-2004 (first entry)
DE Human phosphorylated ERK protein, target for HER-2 therapy SeqID 3.
XX human; HER-2 directed therapy; tumour; insulin growth factor receptor;
KW EGFR; epidermal growth factor receptor; EGFR; S6 ribosomal protein; AKT;
KW NDF; ERK; cancer therapy; predictive biomarker; HER-2/neu.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Modified-site 202 /note= "Optionally a phosphorylated serine (if threonine
FT at 204 is unmodified)"
FT Modified-site 204 /note= "Optionally a phosphorylated threonine (if serine
FT at 202 is unmodified)"
XX WO2004053497-A2.
XX 24-JUN-2004.
XX 11-DEC-2003; 2003WO-US039770.
XX 11-DEC-2002; 2002US-0432942P.
XX (VENT-) VENTANA MEDICAL SYSTEMS INC.
XX (CELL-) CELL SIGNALING TECHNOLOGY INC.
XX Bacus SS, Smith BL;
XX WPI; 2004-507069/48.
XX Identifying a mammalian tumor, useful for predicting a response to HER2-
PT directed therapy, comprises assaying a sample to detect a pattern of

PT	expression, phosphorylation or both, of one or more polypeptides.	
XX	Claim 35; SEQ ID NO 3; 49pp; English.	
XX	This invention relates to a novel method for determining or predicting the response of a patient to HER-2 directed therapy. Specifically, it refers to analysing a mammalian tumour in order to detect a pattern of expression and/or phosphorylation of a protein taken from the group including insulin growth factor receptor (IGFR) polypeptide, epidermal growth factor receptor (EGFR), phosphorylated S6 ribosomal protein, phosphorylated AKT, phosphorylated NDF or phosphorylated ERK protein. The present invention describes characterising a mammalian tumour's responsiveness to an HER-2 therapy and hence an individual's response to this cancer therapy, using an immunologically specific antibody directed against one of the aforementioned proteins. Furthermore, it provides predictive biomarkers that can be used to assess the efficacy of therapeutic agents targeted to HER-2/neu. This polypeptide sequence is the human phosphorylated ERK protein of the invention.	
XX	Sequence 379 AA;	
SQ	Query Match 98.2%; Score 1731; DB 8; Length 379; Best Local Similarity 88.4%; Pred. No. 1.3e-170; Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;	
Qy	1 MAAAAGGGGGPRRTGEGVGPGEVEMVKQPDFVGPRTQLOYIGEGAYGMVSSAY 60 	FT
Db	1 MAAAAGGGGGPRRTGEGVGPGEVEMVKQPDFVGPRTQLOYIGEGAYGMVSSAY 60 	PN
Qy	61 DHVKTTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120 	XX
Db	61 DHVKTTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120 	XX
Qy	121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDL 180 	PR
Db	121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDL 180 	PR
Qy	181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240 	XX
Db	181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240 	XX
Qy	241 NRPIFGKHYLDQNLHIL----- 258 	XX
Db	241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINRKARNYIQLSPSTKVAMAKLFPKSD 300 	XX
Qy	259 --ALDLDRLMTNPNKRITVEEALHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRL 316 	XX
Db	301 SKALDLDRLMTNPNKRITVEEALHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRL 360 	XX
Qy	317 KELIFQETARFQGVLEAP 335 	XX
Db	361 KELIFQETARFQGVLEAP 379 	XX
RESULT 12		
ADV42070	ADV42070 standard; protein; 379 AA.	
XX	ADV42070;	
AC	24-FEB-2005 (first entry)	
DT	Human extracellular signal regulated kinase SEQ ID NO:3.	
XX	tumor; antibody therapy; human epidermal growth factor receptor 2; HER2; phosphorylation; extracellular signal regulated kinase; ERK.	
XX	Homo sapiens.	
OS	Key Location/Qualifiers	
XX	Modified-site 202	
FT	Modified-site 204	
FT	Modified-site 204	

Db 241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINMKARNYLQSLPSKTKVAMAKLPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKRL 316
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKRL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379
RESULT 13
ADW12905
ID ADW12905 standard; protein; 379 AA.
XX
AC ADW12905;
XX
DT 07-APR-2005 (first entry)
XX
DE Human SNAPK3V3 variant protein.
XX
KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW enzyme.
XX
OS Homo sapiens.
XX
PN US2005013817-A1.
XX
PD 20-JAN-2005.
XX
PF 18-JUL-2003; 2003US-00623108.
XX
PR 18-JUL-2003; 2003US-00623108.
XX
PA (DAIK/) DAI K.
XX
PI Dai K;
XX
WPI; 2005-080923/09.
DR N-PSDB; ADW12904.
XX
New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
PT diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
PT in a mammal.
XX
PS Disclosure; SEQ ID NO 6; 55pp; English.
XX
The invention relates to novel isolated polypeptide (I) comprising an
CC amino acid sequence selected from sequences comprising 335 or 359 amino
CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
CC associated with the deficiency of the SNAPK3 gene in a mammal. The
CC fragments of the polypeptides and polynucleotides can also be used as
CC primers or probes. This sequence corresponds to the SNAPK3V1 variant
CC protein.
XX
SQ Sequence 379 AA;
Query Match 98.2%; Score 1731; DB 9; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAGGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGEGYGMVSSAY 60
Db 1 MAAAAAGGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGEGYGMVSSAY 60
QY 61 DHVTRTRVAIKKISPFHQTCQRTLRBIQILLRFRHENVIGIRDILRSTLEAMRDYVI 120
Db 61 DHVTRTRVAIKKISPFHQTCQRTLRBIQILLRFRHENVIGIRDILRSTLEAMRDYVI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQVILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQVILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINMKARNYLQSLPSKTKVAMAKLPKSD 258
Db 241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINMKARNYLQSLPSKTKVAMAKLPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKRL 316
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKRL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379
RESULT 14
ADZ65036
ID ADZ65036 standard; protein; 379 AA.
XX
AC ADZ65036;
XX
DT 14-JUL-2005 (first entry)
XX
DE Mitogen activated protein kinase, MAPK-3, SEQ ID 2.
XX
KW MAP kinase; Mitogen-activated protein kinase; enzyme; cancer.
XX
OS Homo sapiens.
XX
PN US2005095657-A1.
XX
PD 05-MAY-2005.
XX
PF 21-SEP-2004; 2004US-00945684.
XX
PR 11-OCT-2002; 2002US-0418038P.
XX
PR 14-OCT-2003; 2003WO-US032248.
XX
PA (ARBI/) ARBISER J L.
XX
PA (COHE/) COHEN C.
XX
PI Arbiser JL, Cohen C;
XX
WPI; 2005-354474/36.
DR REFSEQ; XP_055766.3.
XX
Detecting phosphorylated mitogen activated protein kinase (P-MAPK), by
PT contacting sample with antibody having affinity for P-MAPK, detecting
PT antibody/P-MAPK complex to detect P-MAPK, and expression of P-MAPK
PT indicates cancer.
XX
PS Claim 3; SEQ ID NO 2; 23pp; English.
XX
The present invention relates to a method (M1) for detecting
CC Phosphorylated Mitogen Activated Protein Kinase (P-MAPK; ADZ65035-
CC ADZ65042), by providing sample, contacting the sample with at least one
CC antibody having affinity for the phosphorylated portion of P-MAPK, and
CC detecting antibody/P-MAPK complex, in which the presence of antibody/P-
CC MAPK complex indicates that P-MAPK is present in sample, and where
CC constitutive expression of P-MAPK is indicative of cancer.
XX
SQ Sequence 379 AA;
Query Match 98.2%; Score 1731; DB 9; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAGGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGEGYGMVSSAY 60
Db 1 MAAAAAGGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGEGYGMVSSAY 60

QY 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
 Db 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
 QY 121 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 Db 121 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
 Db 181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
 QY 241 NRPIFGKHLYDLQNLHIL----- 258
 Db 241 NRPIFGKHLYDLQNLHIL----- 258
 QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTFAMELDDLPKERL 316
 Db 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTFAMELDDLPKERL 360
 QY 317 KELIFOETARFQPGVLEAP 335
 Db 361 KELIFOETARFQPGVLEAP 379

RESULT 15
 AAM40540
 ID AAM40540 standard; protein; 403 AA.
 XX
 AC AAM40540;
 XX
 DT 22-OCT-2001 (first entry)
 DE
 DE Human polypeptide SEQ ID NO 5471.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200153312-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX 26-DEC-2000; 2000WO-US034263.
 PF
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB; AAT59696.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 PT
 XX

PS
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA138642-AA142213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 403 AA;

Query Match 98.2%; Score 1731; DB 4; Length 403;
 Best Local Similarity 88.4%; Pred. No. 1.4e-170; Indels 44; Gaps 1;
 Matches 335; Conservative 0; Mismatches 0;
 QY 1 MAAAAAQQGGGGEPRRTGEGVGFGEVEMVKQPPFDVGPRYTQIQYIGEGAYGMVSSAY 60
 Db 25 MAAAAAQQGGGGEPRRTGEGVGFGEVEMVKQPPFDVGPRYTQIQYIGEGAYGMVSSAY 84
 QY 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
 Db 85 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 144
 QY 121 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 Db 145 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 204
 QY 181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
 Db 205 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 264
 QY 241 NRPIFGKHLYDLQNLHIL----- 258
 Db 265 NRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLSQSPSKTKVAVAKLFPKSD 324
 QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTFAMELDDLPKERL 316
 Db 325 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTFAMELDDLPKERL 384
 QY 317 KELIFOETARFQPGVLEAP 335
 Db 385 KELIFOETARFQPGVLEAP 403

Search completed: February 6, 2006, 15:29:01
 Job time : 274.438 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:29:28 ; Search time 25.1009 Seconds
(without alignments)
1284.124 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAGGGGPRTEGV.....LKELIFQETARFGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1725	97.8	379	1 A48082	MAP kinase 3 (EC 2
2	1656	93.9	380	2 JC1451	Ca2+/calmodulin-de
3	1655	93.9	377	2 S28184	Ca2+/calmodulin-de
4	1636	92.8	369	2 A60041	Ca2+/calmodulin-de
5	1455.5	82.6	392	2 JW0052	extracellular sign
6	1445.5	82.0	360	2 S25011	protein kinase ERK
7	1444.5	81.9	360	1 JQ1400	MAP kinase 1 (EC 2
8	1434.5	81.4	358	2 S16444	mitogen-activated
9	1434.5	81.4	358	2 A40033	protein kinase (EC
10	1433	81.3	369	2 JW0053	extracellular sign
11	1425	80.8	361	2 A39754	mitogen-activated
12	1293.5	73.4	376	2 A46036	extracellular sign
13	1272	72.1	376	2 A36978	MAP kinase mpk-1 (
14	1272	72.1	444	2 A36977	MAP kinase sur-1 (
15	907	51.4	356	2 T51944	pathogenicity MAP
16	883.5	50.1	361	2 T51943	mitogen-activated
17	855.5	48.5	362	2 A47211	protein kinase ERK
18	844	47.9	415	2 A56042	mitogen-activated
19	840	47.6	372	2 S15663	protein kinase (EC
20	833.5	47.3	387	2 S48123	mitogen-activated
21	828	47.0	394	2 S33635	mitogen-activated
22	821	46.6	393	2 S51321	mitogen-activated
23	820	46.5	395	2 S40472	mitogen-activated
24	808.5	45.9	371	2 S51320	mitogen-activated
25	804	45.6	393	2 T47803	mitogen-activated
26	803.5	45.6	370	2 T47504	mitogen-activated
27	803.5	45.6	370	2 S40469	mitogen-activated
28	802	45.5	376	2 S40470	mitogen-activated
29	801.5	45.5	371	2 T09622	protein kinase MMK

30	799.5	45.3	369	2 S56638	mitogen-activated
31	792.5	45.0	371	2 T14915	mitogen-activated
32	792.5	45.0	376	2 S40471	mitogen-activated
33	781	44.3	372	2 S39559	mitogen-activated
34	781	44.3	375	2 T03971	mitogen-activated
35	781	44.3	384	2 S52889	mitogen-activated,
36	780	44.2	376	2 F96619	protein T30E16.13
37	779.5	44.2	371	2 S60121	mitogen-activated
38	779.5	44.2	815	2 B56708	extracellular sign
39	769.5	43.6	363	2 C86214	hypothetical prote
40	769	43.6	368	2 A33297	probable serine/th
41	768	43.6	370	2 F85236	protein F14N23.9
42	767.5	43.5	369	2 C86146	hypothetical prote
43	761	43.2	353	2 S28548	protein kinase (EC
44	761	43.2	373	2 T13024	probable protein k
45	760.5	43.1	349	2 S68675	mitogen-activated

ALIGNMENTS

RESULT 1

A48082

MAP kinase 3 (EC 2.7.1.-) - human

N;Alternate names: extracellular signal-regulated kinase 1 (ERK1); mitogen-activated prot

N;Contains: protein kinase (EC 2.7.1.37)

C;Species: Homo sapiens (man)

C;Date: 04-Sep-1998 #sequence revision 04-Sep-1998 #text_change 09-Jul-2004

C;Accession: A48082; PQ0270; S23428; S15519; S21579

R;Charest, D.L.; Mordret, G.; Harder, K.W.; Jirik, F.; Pelech, S.L.

Mol. Cell. Biol. 13, 4679-4690, 1993

A;Title: Molecular cloning, expression, and characterization of the human mitogen-activat

A;Reference number: A48082; MUID:93330262; PMID:7687743

A;Accession: A48082

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-379 <CHA>

A;Cross-references: UNIPROT:P27361; UNIPARC:UPI0000035BE2; EMBL:X60188; NID:g31220; PIDN

A;Experimental source: hepatoma cell line HEP G2

R;Note: authors translated the codon AGC for residue 174 as Ile

R;Owaki, H.; Makar, R.; Boulton, T.G.; Cobb, M.H.; Geppert, T.D.

Biochem. Biophys. Res. Commun. 182, 1416-1422, 1992

A;Title: Extracellular signal-regulated kinases in T cells: characterization of human ERK

A;Reference number: JQ1400; MUID:92171961; PMID:1540184

A;Accession: PQ0270

A;Molecule type: mRNA

A;Residues: 14-173, 'I', 175-379 <OWA>

A;Cross-references: UNIPARC:UPI000016AB99; GB:M84490; NID:g186695; PIDN:AAA36142.1; PID:9

R;Experimental source: cell line CEM

R;Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.

FEBS Lett. 304, 170-178, 1992

A;Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.

A;Reference number: S23426; MUID:92316223; PMID:1319925

A;Accession: S23428

A;Molecule type: mRNA

A;Residues: 25-373, 'I', 175-379 <CON>

A;Cross-references: UNIPARC:UPI000016A2B4; EMBL:Z11696; NID:g23882; PIDN:CAA7754.1; PID

C;Comment: This enzyme is activated by MAP kinase kinase (see PIR:45100 and PIR:A46723).

C;Genetics:

A;Gene: GDB:PRK3; ERK1

A;Cross-references: GDB:135679; OMIM:601795

A;Map position: 16pter-16qter

C;Complex: monomer

C;Function:

A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine

A;Pathway: MAP kinase cascade

C;Superfamily: kinase-related

C;Keywords: ATP; monomer; phosphoprotein; phosphotransferase; serine/threonine-specific f

F;40-330/Domain: protein kinase homology <KIN>

F;48-56/Region: protein kinase ATP-binding motif

F;202/Binding site: phosphate (Thr) (covalent) (by MAP kinase kinase) #status predicted

F;204/Binding site: phosphate (Tyr) (covalent) (by MAP kinase kinase) #status predicted

Query Match		97.8%;	Score 1725;	DB 1;	Length 379;
Best Local Similarity		88.1%;	Pred. No. 1.8e-70;		
Matches 334;		Conservative 0;	Mismatches 1;	Indels 44;	Gaps 1;
QY	1	MAAAAQGGGGGPRTEGVGPVGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAY	60		
Db	1	MAAAAQGGGGGPRTEGVGPVGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAY	60		
QY	61	DHVKTRVAIKKISPEHOTYCORTLREIQILLRPHENVIGIRDILRASTLEAMRDVVI	120		
Db	61	DHVKTRVAIKKISPEHOTYCORTLREIQILLRPHENVIGIRDILRASTLEAMRDVVI	120		
QY	121	VDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVHLRDLKPSNLLINTTCDL	180		
Db	121	VDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVHLRDLKPSNLLINTTCDL	180		
QY	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDISVSGCILAEMLS	240		
Db	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDISVSGCILAEMLS	240		
QY	241	NRPIFGKHVLDOLNHIL-----	258		
Db	241	NRPIFGKHVLDOLNHILGILGSPSEDNLNCIINKARNYLQSLPSKTKVAVAKLPFKSD	300		
QY	259	--ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPERLK	316		
Db	301	SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPERLK	360		
QY	317	KELIFQETARFQGVLEAP 335			
Db	361	KELIFQETARFQGVLEAP 379			
RESULT 2					
JC1451					
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) - rat					
N/Alternate names: ERK1-MAP kinase; extracellular signal-regulated kinase					
C/Species: Rattus norvegicus (Norway rat)					
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004					
C/Accession: JC1451; A35061; A37140; A40466; S24947					
R/Marquardt, B.; Stabel, S.					
Gene 120, 297-299, 1992					
A/Title: Sequence of a rat cDNA encoding the ERK1-MAP kinase.					
A/Reference number: JC1451; MUID:93013050; PMID:1327976					
A/Accession: JC1451					
A/Molecule type: mRNA					
A/Residues: 1-380 <WAR>					
A/Cross-references: UNIPROT:P21708; UNIPARC:UPI000012F174; EMBL:X65198; NID:g56626; PIDN					
A/Experimental source: brain					
R/Boulton, T.G.; Yancopoulos, G.D.; Gregory, J.S.; Slaughter, C.; Moomaw, C.; Hsu, J.; C					
Science 249, 64-67, 1990					
A/Title: An insulin-stimulated protein kinase similar to yeast kinases involved in cell					
A/Reference number: A35061; MUID:90312137; PMID:2164259					
A/Accession: A35061					
A/Molecule type: mRNA					
A/Residues: 14-380 <BOU>					
A/Cross-references: UNIPARC:UPI0000145072; GB:M38194; NID:g204051; PIDN:AAA41123.1; PID:					
R/Boulton, T.G.; Gregory, J.S.; Cobb, M.H.					
Biochemistry 30, 278-286, 1991					
A/Title: Purification and properties of extracellular signal-regulated kinase 1, an insu					
A/Reference number: A37140; MUID:91105092; PMID:1846291					
A/Accession: A37140					
A/Status: preliminary					
A/Molecule type: protein					
A/Residues: 43-64;167-178, 'X',180-183, 'X',185 <BO2>					
A/Cross-references: UNIPARC:UPI000017558F; UNIPARC:UPI0000175590					
R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.					
DNA Cell Biol. 10, 505-514, 1991					
A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse a					
A/Reference number: A40466; MUID:91369479; PMID:1716439					
A/Accession: A40466					
A/Status: preliminary					
A/Molecule type: mRNA					

A/Residues: 14-94, 'R', 96-380 <DEM>									
A/Cross-references: UNIPARC:UPI00001707CE; GB:U12008; GB:S59509; NID:g515498; PIDN:AAA200									
C/Superfamily: kinase-related transforming protein; protein kinase homology									
C/Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; serine/threonine									
F:41-331/Domain: protein kinase homology <KIN>									
F:49-57/Region: protein kinase ATP-binding motif									
Query Match 93.9%; Score 1656; DB 2; Length 380;									
Best Local Similarity 85.4%; Pred. NO. 2.1e-67;									
Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;									
QY	2	AAAAAQQGGGGPRTEGVGPVGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAYD	61						
DB	3	AAAAAAPGGGGPRGTAGVVPVPGVEVVRKQPPDVGPRTYQLOYIGEGAYGMVSSAYD	62						
QY	62	HVKTRVAIKKISPEHOTYCORTLREIQILLRPHENVIGIRDILRASTLEAMRDVYIV	121						
DB	63	HVKTRVAIKKISPEHOTYCORTLREIQILLRPHENVIGIRDILRASTLEAMRDVYIV	122						
QY	122	QDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVHLRDLKPSNLLINTTCDLK	181						
DB	123	QDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVHLRDLKPSNLLINTTCDLK	182						
QY	182	ICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDISVSGCILAEMLSN	241						
DB	183	ICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDISVSGCILAEMLSN	242						
QY	242	RFIPFGKHVLDQLNHIL-----	258						
DB	243	RFIPFGKHVLDQLNHILGILGSPQEDNLNCIINKARNYLQSLPSKTKVAVAKLPFKSDS	302						
QY	259	-ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPERLK	317						
DB	303	KALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPERLK	362						
QY	318	ELIFQETARFQGVLEAP	335						
DB	363	ELIFQETARFQGAPEAP	380						
RESULT 3									
S28184									
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) erk-1 - mouse (fragment)									
N/Alternate names: extracellular signal-regulated kinase 1									
C/Species: Mus musculus (house mouse)									
C/Date: 22-Nov-1993 #sequence_revision 30-Jan-1998 #text_change 09-Jul-2004									
C/Accession: S28184; B40466; A41371									
R/Tanner, B.; Mueckler, M.									
Biochim. Biophys. Acta 1171, 319-320, 1993									
A/Title: Molecular cloning of a mouse extracellular signal regulated kinase (erk-1).									
A/Reference number: S28184; MUID:93144347; PMID:8424957									
A/Accession: S28184									
A/Molecule type: mRNA									
A/Residues: 1-377 <TAN>									
A/Cross-references: UNIPROT:Q63844; UNIPARC:UPI000017558C; EMBL:Z14249									
R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.									
DNA Cell Biol. 10, 505-514, 1991									
A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse									
A/Reference number: A40466; MUID:91369479; PMID:1716439									
A/Accession: B40466									
A/Molecule type: mRNA									
A/Residues: 74-377 <DEM>									
A/Cross-references: UNIPARC:UPI0000023599; GB:S59517									
R/Crews, C.M.; Alessandrini, A.A.; Erikson, R.L.									
Proc. Natl. Acad. Sci. U.S.A. 88, 8845-8849, 1991									
A/Title: Mouse Erk-1 gene product is a serine/threonine protein kinase that has the pote									
A/Reference number: A41371; MUID:92020947; PMID:1717989									
A/Accession: A41371									
A/Molecule type: mRNA									
A/Residues: 7-16 <CRE>									
A/Cross-references: UNIPARC:UPI0000170C57; GB:S58470; NID:g236372; PIDN:AAB19973.1; PID:Q									
C/Superfamily: kinase-related transforming protein; protein kinase homology									
C/Keywords: ATP; calmodulin binding; phosphotransferase; protein kinase									

db 121 KLKSOOLSNPDHICYFLYOILRGLKYIHSANVLHRDLKPSNLLINTTCDIKICDFGLARI 18

RESORT 9
S25011

RESULT 6
S25011

protein kinase ERK2 (EC 2.7.1.-) - bovine
N:Alternate names: extracellular signal-regulated kinase
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S25011
R:Ely, C.M.; Cox, M.E.; Her, J.; Parsons, S.J.
submitted to the EMBL Data Library, July 1992
A:Description: Cloning and sequencing of ERK2 from a bovine adrenal medulla cDNA library
A:Reference number: S25011
A:Accession: S25011
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-360 <ELY>
A:Cross-references: UNIPROT:P46196; UNIPARC:UPI000012F172; EMBL:Z14089; NID:g182190; PIDN:
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:23-313/Domain: protein kinase homology <KIN>
F:31-39/Region: protein kinase ATP-binding motif
Query Match 82.0%; Score 1445.5; DB 2; Length 360;
Best Local Similarity 74.8%; Pred. No. 4.6e-58;
Matches 279; Conservative 17; Mismatches 17; Indels 61; Gaps 2;
QY 1 MAAAAAQQGGGPRRTGPGVGPGEVMVGQPPDVGPRYTQIYIGEGAYGMVSSAY 60
DB 1 MAAAAAAGAG-----PEWVRGQVFDVGPRYTNLSYIGEGAYGMVCSAY 43
QY 61 DHVTRKTRVAIKKISPEHQTQRTLEIQTILRRHENHVGIRDIRASTLEAMRDVYI 120
DB 44 DNVNKRVAIKKISPEHQTQRTLEIKILLRFRHENIIGINDIRAPTEIQMKDVI 103
QY 121 VDLMETDLYKLLKSQOLSNHHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 180
DB 104 VDLMETDLYKLLKTOHLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 163
QY 181 KICDFGLARIADPEHDHGTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
DB 164 KICDFGLARVADPDHDHGTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 223
QY 241 NRPIFGKHYLDQNLHIL-----PEWVRGQVFDVGPRYTNLSYIGEGAYGMVSSAY 258
DB 224 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINLKARNYLLSLPHKNKVPWNRLFPNAD 283
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 316
DB 284 SKALDLDLDMKLTFFNPKRIEVEQALAHPLYEQYDPSDEPIAEAFKFDMLDLDLPKEKL 343
RESULT 7
QJ1400
MAP kinase 1 (EC 2.7.1.-) - human
N:Alternate names: ERK; extracellular signal-regulated kinase 2 (ERK2); mitogen-activate
N:Contains: protein kinase (EC 2.7.1.37)
C:Species: Homo sapiens (man)
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
C:Accession: QJ1400; S23426; S23427; S21577; S21578
R:Owaki, H.; Makar, R.; Boulton, T.G.; Cobb, M.H.; Gerecht, T.D.
Biochem. Biophys. Res. Commun. 182, 1416-1422, 1992
A:Title: Extracellular signal-regulated kinases in T cells: characterization of human ER
A:Reference number: QJ1400; MUID:92171961; PMID:1540184
A:Accession: QJ1400
A:Molecule type: mRNA
A:Residues: 1-360 <OWA>
A:Cross-references: UNIPROT:P28482; UNIPARC:UPI000000104P; GB:M84489; NID:g182190; PIDN:
A:Experimental source: cell line Jurkat
R:Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.
FEBS Lett. 304, 170-178, 1992
A:Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.
A:Reference number: S23426; MUID:92316223; PMID:1319925

A:Accession: S23426
A:Molecule type: mRNA
A:Residues: 13-360 <GON1>
A:Cross-references: UNIPARC:UPI000016A28D; EMBL:Z11695; NID:g23878; PIDN:CAA77753.1; PID:
A:Accession: S23427
A:Molecule type: mRNA
A:Residues: 1-90, 'O', '92-360 <GON2>
A:Cross-references: UNIPARC:UPI000016A299; EMBL:Z11694; NID:g23880; PIDN:CAA77752.1; PID:
C:Comment: This enzyme is activated by MAP kinase kinase (see PIR:A46723).
C:Genetics:
A:Gene: GDB:PRXM1; ERK; MAPK1; ERK2
A:Cross-references: GDB:135677; OMIM:176948
A:Map position: 22q11.2-22q11.2
C:Complex: monomer
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
A:Pathway: MAP kinase cascade
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F:23-313/Domain: protein kinase homology <KIN>
F:31-39/Region: protein kinase ATP-binding motif
F:185/Binding site: phosphate (Thr) (covalent) (by MAP kinase kinase) #status predicted
F:187/Binding site: phosphate (Tyr) (covalent) (by MAP kinase kinase) #status predicted
Query Match 81.9%; Score 1444.5; DB 1; Length 360;
Best Local Similarity 74.3%; Pred. No. 5.1e-58;
Matches 278; Conservative 18; Mismatches 17; Indels 61; Gaps 2;
QY 1 MAAAAAQQGGGPRRTGPGVGPGEVMVGQPPDVGPRYTQIYIGEGAYGMVSSAY 60
DB 1 MAAAAAAGAG-----PEWVRGQVFDVGPRYTNLSYIGEGAYGMVCSAY 43
QY 61 DHVTRKTRVAIKKISPEHQTQRTLEIQTILRRHENHVGIRDIRASTLEAMRDVYI 120
DB 44 DNVNKRVAIKKISPEHQTQRTLEIKILLRFRHENIIGINDIRAPTEIQMKDVI 103
QY 121 VDLMETDLYKLLKSQOLSNHHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 180
DB 104 VDLMETDLYKLLKTOHLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 163
QY 181 KICDFGLARIADPEHDHGTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
DB 164 KICDFGLARVADPDHDHGTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 223
QY 241 NRPIFGKHYLDQNLHIL-----PEWVRGQVFDVGPRYTNLSYIGEGAYGMVSSAY 258
DB 224 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINLKARNYLLSLPHKNKVPWNRLFPNAD 283
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 316
DB 284 SKALDLDLDMKLTFFNPKRIEVEQALAHPLYEQYDPSDEPIAEAFKFDMLDLDLPKEKL 343
RESULT 8
S16444
Mitogen-activated protein kinase (EC 2.7.1.-) - mouse
N:Alternate names: cell division cycle-2-related protein kinase ERK2
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: S16444; S15015; FN0480
R:Her, J.H.; Wu, J.; Rall, T.B.; Sturgill, T.W.; Weber, M.J.
Nucleic Acids Res. 19, 3743, 1991
A:Title: Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by tyrosine ph
A:Reference number: S16444; MUID:91305126; PMID:1649458
A:Accession: S16444
A:Molecule type: mRNA
A:Residues: 1-358 <HER>
A:Cross-references: UNIPROT:P27703; UNIPARC:UPI0000003F8A; EMBL:X58712; NID:g53001; PIDN:
A:Note: part of this sequence was confirmed by protein sequencing

R;Payne, D.M.; Rossomando, A.J.; Martino, P.; Erickson, A.K.; Her, J.H.; Shabanowitz, J.
EMBO J. 10, 885-892, 1991
A;Title: Identification of the regulatory phosphorylation sites in pp42/mitogen-activated
A;Reference number: S15015; MUID:91184134; PMID:1849075
A;Accession: S15015
A;Status: preliminary
A;Molecule type: protein
A;Residues: 171-181, 'X', 183-189 <P>A>
A;Cross-references: UNIPARC:UPI0000175588
R;Ershler, M.A.; Nagorekaya, T.V.; Visser, J.W.M.; Belyavsky, A.V.
Gene 124, 305-306, 1993
A;Title: Novel CDC2-related protein kinases produced in murine hematopoietic stem cells.
A;Reference number: PN0479; MUID:93185941; PMID:8444355
A;Accession: PN0480
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 145-195 <RS>
A;Cross-references: UNIPARC:UPI0000175589
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F;21-311/Domain: protein kinase homology <KIN>
F;29-37/Region: protein kinase ATP-binding motif
F;183/Binding site: phosphate (Thr) (covalent) #status experimental
F;185/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 81.4%; Score 1434.5; DB 2; Length 358;
Best Local Similarity 73.8%; Pred. No. 1.4e-57;
Matches 276; Conservative 19; Mismatches 16; Indels 63; Gaps 2;

QY 1 MAAAAAGGGGGEPRRTGEGVGPGEVEMVKGQPFVDGPRYTQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAAG-----PEMVRGQVFDGPRYTNSLYIGEGAYGMVCSAY 41
QY 61 DHVRKTRVAIKKISPFPHQTYCORTLEIQLLRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 42 DNLNKRVAIKKISPFPHQTYCORTLEIKILLFRHENIIGINDIIRAPTIQOMKDVYI 101
QY 121 VQDLMETDLYKLLKSQQLSNDHICFYLYQILRLGYIHSANVLRDLKPSNLLNTTCDL 180
DB 102 VQDLMETDLYKLLKTQHLSDNDHICFYLYQILRLGYIHSANVLRDLKPSNLLNTTCDL 161
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 162 KICDFGLARVADPDHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 221
QY 241 NRPIFGKHYLDQLNHIL----- 258
DB 222 NRPIFGKHYLDQLNHILGILGSPSQEDLNCIINLKARNYLLSLPHKNKVPWNRLFPNAD 281
QY 259 --ALDLDRLMTFNPKNRTVEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 316
DB 282 SKALDLDLKMFTNPKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKEKL 341
QY 317 KELIFOETARFQPG 330
DB 342 KELIFEETARFQPG 355

RESULT 10
JW0053
extracellular signal-regulated kinase (EC 2.7.-.-) 2 - common carp
N;Alternate names: CERK2
C;Species: Cyprinus carpio (common carp)
C;Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C;Accession: JW0053
R;Hashimoto, H.; Yokoyama, Y.; Matsuo, Y.; Toyohara, H.; Kohno, M.; Sakaguchi, M.
J. Biochem. 123, 1031-1035, 1998
A;Title: Existence of two isoforms of extracellular signal-regulated kinase in fish.
A;Reference number: JW0052; MUID:98269030; PMID:9603989
A;Accession: JW0053
A;Molecule type: mRNA
A;Residues: 1-369 <HAS>
A;Cross-references: UNIPROT:Q7LZH5; UNIPARC:UPI0000175597
C;Comment: This enzyme is the key enzyme in the intracellular signal transduction pathway
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase
F;32-322/Domain: protein kinase homology <KIN>

Query Match 81.3%; Score 1433; DB 2; Length 369;
Best Local Similarity 74.3%; Pred. No. 1.7e-57;
Matches 277; Conservative 17; Mismatches 25; Indels 54; Gaps 3;

QY 2 AAAAAAGGGGGEPRRTGEGVGPGEVEMVKGQPFVDGPRYTQIQYIGEGAYGMVSSAYD 61
DB 4 AAVSAPAGGGPNP-----GSG-----AEVVRGQAFDVGPRYTNSLYIGEGAYGMVCSAYK 53
QY 62 HVRKTRVAIKKISPFPHQTYCORTLEIQLLRFRHENVIGIRDLRASTLEAMRDVYIV 121
DB 54 RDNKRVVAIKKISPFPHQTYCORTLEIKILLFRKENIIGINDIIRTTTIDQMKDVYIV 113
QY 122 QDLMETDLYKLLKSQQLSNDHICFYLYQILRLGYIHSANVLRDLKPSNLLNTTCDLK 181
DB 114 QDLMETDLYKLLKTQHLSDNDHICFYLYQILRLGYIHSANVLRDLKPSNLLNTTCDLK 173

R;Payne, D.M.; Rossomando, A.J.; Martino, P.; Erickson, A.K.; Her, J.H.; Shabanowitz, J.
EMBO J. 10, 885-892, 1991
A;Title: Identification of the regulatory phosphorylation sites in pp42/mitogen-activated
A;Reference number: S15015; MUID:91184134; PMID:1849075
A;Accession: S15015
A;Status: preliminary
A;Molecule type: protein
A;Residues: 171-181, 'X', 183-189 <P>A>
A;Cross-references: UNIPARC:UPI0000175588
R;Ershler, M.A.; Nagorekaya, T.V.; Visser, J.W.M.; Belyavsky, A.V.
Gene 124, 305-306, 1993
A;Title: Novel CDC2-related protein kinases produced in murine hematopoietic stem cells.
A;Reference number: PN0479; MUID:93185941; PMID:8444355
A;Accession: PN0480
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 145-195 <RS>
A;Cross-references: UNIPARC:UPI0000175589
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F;21-311/Domain: protein kinase homology <KIN>
F;29-37/Region: protein kinase ATP-binding motif
F;183/Binding site: phosphate (Thr) (covalent) #status experimental
F;185/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 81.4%; Score 1434.5; DB 2; Length 358;
Best Local Similarity 73.8%; Pred. No. 1.4e-57;
Matches 276; Conservative 19; Mismatches 16; Indels 63; Gaps 2;

QY 1 MAAAAAGGGGGEPRRTGEGVGPGEVEMVKGQPFVDGPRYTQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAAG-----PEMVRGQVFDGPRYTNSLYIGEGAYGMVCSAY 41
QY 61 DHVRKTRVAIKKISPFPHQTYCORTLEIQLLRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 42 DNLNKRVAIKKISPFPHQTYCORTLEIKILLFRHENIIGINDIIRAPTIQOMKDVYI 101
QY 121 VQDLMETDLYKLLKSQQLSNDHICFYLYQILRLGYIHSANVLRDLKPSNLLNTTCDL 180
DB 102 VQDLMETDLYKLLKTQHLSDNDHICFYLYQILRLGYIHSANVLRDLKPSNLLNTTCDL 161
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 162 KICDFGLARVADPDHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 221
QY 241 NRPIFGKHYLDQLNHIL----- 258
DB 222 NRPIFGKHYLDQLNHILGILGSPSQEDLNCIINLKARNYLLSLPHKNKVPWNRLFPNAD 281
QY 259 --ALDLDRLMTFNPKNRTVEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 316
DB 282 SKALDLDLKMFTNPKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKEKL 341
QY 317 KELIFOETARFQPG 330
DB 342 KELIFEETARFQPG 355

RESULT 9
A40033
protein kinase (EC 2.7.1.37) ERK2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004
C;Accession: A40033
R;Boulton, T.G.; Nye, S.H.; Robbins, D.J.; Ip, N.Y.; Radziejewska, E.; Morgenbesser, S.D.
Cell 65, 663-675, 1991
A;Title: ERKs: a family of protein-serine/threonine kinases that are activated and tyros
A;Reference number: A40033; MUID:91235302; PMID:2032290
A;Accession: A40033
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <BOU>
A;Cross-references: UNIPROT:P27703; UNIPARC:UPI00000003FEA; GB:M64300; NID:g204055; PIDN:

```
QY 182 ICDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 241
Db 174 ICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 233
QY 242 RPIFFPGKHVLDQLNHIL----- 258
Db 234 RPIFFPGKHVLDQLNHILGILGSPSQEDLNLCINIKARNYLLSLPLCKVPNRLFPNADP 293
QY 259 -ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKERUK 317
Db 294 KALDLLDKMLTFNPHKRIVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKETLK 353
QY 318 ELIFQETARFQPG 330
Db 354 ELIFEETARFQPG 366

RESULT 11
A39754
mitogen-activated protein kinase (EC 2.7.1.1) - African clawed frog
N:Alternate names: M phase MAP kinase
C:Species: Xenopus laevis (African clawed frog)
C:Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 09-Jul-2004
C:Accession: A39754; S16597
R:Posada, J.; Sanghera, J.; Palech, S.; Aebersold, R.; Cooper, J.A.
Mol. Cell. Biol. 11, 2517-2528, 1991
A:Title: Tyrosine phosphorylation and activation of homologous protein kinases during oocyte maturation
A:Reference number: A39754; MUID:91203872; PMID:1708093
A:Accession: A39754
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-361 <POS>
A:Cross-references: UNIPROT:P26696; UNIPARC:UPI0000175592; GB:M60977
R:Gotoh, Y.; Moriyama, K.; Matsuda, S.; Okumura, E.; Kishimoto, T.; Kawasaki, H.; Suzuki, EMO J. 10, 2661-2668, 1991
A:Title: Xenopus M phase MAP kinase: isolation of its cDNA and activation by MPF.
A:Reference number: S16597; MUID:91330892; PMID:1714387
A:Accession: S16597
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4,'G','6','A','8-28','1','30-31','A','33-46','D','48','V','50-331','F',333-361 <GOT>
A:Cross-references: UNIPARC:UPI000017156A; EMBL:X59813; NID:G64893; PIDN:CAA42482.1; PID:CAA42482.1
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:26-316/Domain: protein kinase homology <KIN>
F:34-42/Region: protein kinase ATP-binding motif

Query Match 80.8%; Score 1425; DB 2; Length 361;
Best Local Similarity 73.5%; Pred. No. 3.7e-57;
Matches 275; Conservative 18; Mismatches 23; Indels 58; Gaps 2;

QY 1 MAAAAAGGGGGPRRTGPGVGPGEVEMVKGQPFVGPVGTQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAASSNPGGSP-----EMVRQAQFDVGPVYTNLSYIGEGAYGMVCSAH 46
QY 61 DHVKTQRTVAIKKISPEHQTYCQRTIREIQILLRPHENVIGIRDLIRASTLEAMRDVYI 120
Db 47 CNINKVRVAIKKISPEHQTYCQRTIREIKILLRPHENIIGINDIRAPTIEQMKDVYI 106
QY 121 VQDLMTDLYKLLKSQOLSNHICFYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDL 180
Db 107 VQDLMTDLYKLLKTOHLSNDHICFYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDL 166
QY 181 KICDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 240
Db 167 KICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 226
QY 241 NRPIFFPGKHVLDQLNHIL----- 258
Db 227 NRPIFFPGKHVLDQLNHILGILGSPSQEDLNLCINILKARNYLLSLPHKNKVPNRLFPNAD 286
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKERL 316
```

```
Db 287 PKALDLLDKMLTFNPHKRIVEEALAHPLYEQYYDPSPEVAEAPLKFEMELDDLPKETL 346
QY 317 KELIFQETARFQPG 330
Db 347 KELIFEETARFQPG 360

RESULT 12
A46036
extracellular signal-regulated kinase (EC 2.7.1.1) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46036
R:Biggs III, W.H.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 89, 6295-6299, 1992
A:Title: Primary structure, expression, and signal-dependent tyrosine phosphorylation of the Drosophila melanogaster extracellular signal-regulated kinase
A:Reference number: A46036; MUID:92335284; PMID:1378625
A:Accession: A46036
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-376 <BIG>
A:Cross-references: UNIPROT:P40417; UNIPARC:UPI0000000PB3; GB:M95124; NID:G485754; PIDN:G485754
A:Note: sequence extracted from NCBI backbone (NCBI:108684, NCBI:108685)
C:Genetics:
A:Gene: FlyBase:rl
A:Cross-references: FlyBase:FBgn0003256
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:36-326/Domain: protein kinase homology <KIN>
F:44-52/Region: protein kinase ATP-binding motif

Query Match 73.4%; Score 1293.5; DB 2; Length 376;
Best Local Similarity 71.7%; Pred. No. 2.7e-51;
Matches 251; Conservative 22; Mismatches 32; Indels 45; Gaps 2;

QY 24 VP-GEVEMVKGQPFVGPVGTQLOVIGEGAYGMVSSAYDHVKTQRTVAIKKISPEHQTYC 82
Db 19 VPQSNAEVIRGQIFEVGPVYIKLAVIGEGAYGMVVSADDTLLNQRTVAIKKISPEHQTYC 78
QY 83 QRTLEIQILLRPHENVIGIRDLIRASTLEAMRDVYIVQDLMTDLYKLLKSQOLSNH 142
Db 79 QRTLEITILTRFKHENIIDIRDLIRVDSIDQMRDVYIVQCLMETDLYKLLKTQRLSNH 138
QY 143 ICYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDLKICDFGLARIADPHDHTGFLT 202
Db 139 ICYFLYQILRLKVIHNSANVLRDLKPSNLLNKTCDLKICDFGLARIADPHDHTGFLT 198
QY 203 EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHVLDQLNHIL 257
Db 199 EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHVLDQLNHILGVLG 258
QY 258 -----LALDLLDRMLTFNPNKRITVE 278
Db 259 SPSRDDLECIINEKARNYLLSLPFPNVPWAKLFPNADALADLLGKMLTFNPHKRIPVE 318
QY 279 EALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKERLKEIFQETARFQ 328
Db 319 EALAHPLYEQYYDGPDEPVAEVFPFRINMENDDISDALKSLIFEETLKF 368

RESULT 13
A36978
MAP kinase mpk-1 (EC 2.7.1.1) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A36978; T22118
R:Lackner, M.R.; Kornfeld, K.; Miller, L.M.; Horvitz, H.R.; Kim, S.K.
Genes Dev. 8, 160-173, 1994
A:Title: A MAP kinase homolog, mpk-1, is involved in ras-mediated induction of vulval cell division
A:Reference number: A36978; MUID:94131270; PMID:8299936
A:Accession: A36978
A:Status: preliminary
```

A:Molecule type: mRNA
A:Residues: 1-376 <LAC>
A:Cross-references: UNIPROT:P39745; UNIPARC:UPI000002B465; GB:U271124; GB:S68854; NID:986
R:Jassal, B.; Smith, A.
Submitted to the EMBL Data Library, December 1994
A:Reference number: Z19519
A:Accession: T22118
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-376 <WTL>
A:Cross-references: UNIPARC:UPI000002B465; EMBL:Z46937; PIDN:CAA87057.1; GSPDB:GN000021;
A:Experimental source: clone F43C1
C:Genetics:
A:Gene: mpk-1
A:Map position: 3
A:Introns: 23/3; 43/2; 85/3; 104/2; 304/2
A:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F;26-285/Domain: protein kinase homology <KIN>
F;34-42/Region: protein kinase ATP-binding motif

Query Match 72.1%; Score 1272; DB 2; Length 376;
Best Local Similarity 70.0%; Pred. No. 2.5e-50;
Matches 238; Conservative 33; Mismatches 25; Indels 44; Gaps 1;

Qy 28 VEMVKGQFPDVGPRYTQLOYIGEGAYGVSSAYDHYVTRVAKKISPFPHQTYCQRTL 87
Db 14 VEEVHGQLFEVAPRYVNLVIGEGAYGVASALDTITRDRVAIKKISPFPHQTYCQRTL 73
Qy 88 EIQILLRPHENVIGIRDIRASTLEAMRDVYIVQDLMETDLYKLKSKQSLNDHCYFL 147
Db 74 EIKILNRFKHENIINQIETSRSETVDSKDIYVQCLMETDLYKLKSKQSLNDHCYFL 133
Qy 148 YQILRGKLYTHSANVLRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGFLTEYVAT 207
Db 134 YQILRGKLYTHSANVLRDLKPSNLLNTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 193
Qy 208 RWTYRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPFPGKHLYDQNLHIL----- 258
Db 194 RWTYRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPFPGKHLYDQNLHILAVVGSNSNA 253
Qy 259 -----ALDILDRMLTFNPKRITVEEALAH 293
Db 254 DLQCIINDKARSYLIISLPHKPKQPWALYFGADPRALDLDKMLTFNPHNRIDIEQALAH 313
Qy 284 PYLEQYVDPTDPRVAEPEPTFAMELDDLPKRLKELIFOE 323
Db 314 PYLEQYDPGDEPVCCEPFTLEMEFDDLPKRLKELIWE 353

RESULT 14
A36977
MAP kinase sur-1 (EC 2.7.1.1) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A36977
R:Wu, Y.; Han, M.
Genes Dev. 8, 147-159, 1994
A:Title: Suppression of activated Let-60 ras protein defines a role of Caenorhabditis el
A:Reference number: A36977; MUID:94131269; PMID:8299935
A:Accession: A36977
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-444 <WUA>
A:Cross-references: UNIPROT:P39745; UNIPARC:UPI0000164181; GB:U03879; NID:G433356; PIDN:
C:Genetics:
A:Gene: sur-1
A:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F;94-353/Domain: protein kinase homology <KIN>
F;102-110/Region: protein kinase ATP-binding motif

Query Match 72.1%; Score 1272; DB 2; Length 444;

Best Local Similarity 70.0%; Pred. No. 2.9e-50;
Matches 238; Conservative 33; Mismatches 25; Indels 44; Gaps 1;

Qy 28 VEMVKGQFPDVGPRYTQLOYIGEGAYGVSSAYDHYVTRVAKKISPFPHQTYCQRTL 87
Db 82 VEEVHGQLFEVAPRYVNLVIGEGAYGVASALDTITRDRVAIKKISPFPHQTYCQRTL 141
Qy 88 EIQILLRPHENVIGIRDIRASTLEAMRDVYIVQDLMETDLYKLKSKQSLNDHCYFL 147
Db 142 EIKILNRFKHENIINQIETSRSETVDSKDIYVQCLMETDLYKLKSKQSLNDHCYFL 201
Qy 148 YQILRGKLYTHSANVLRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGFLTEYVAT 207
Db 202 YQILRGKLYTHSANVLRDLKPSNLLNTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 261
Qy 208 RWTYRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPFPGKHLYDQNLHIL----- 258
Db 262 RWTYRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPFPGKHLYDQNLHILAVVGSNSNA 321
Qy 259 -----ALDILDRMLTFNPKRITVEEALAH 293
Db 322 DLQCIINDKARSYLIISLPHKPKQPWALYFGADPRALDLDKMLTFNPHNRIDIEQALAH 381
Qy 284 PYLEQYVDPTDPRVAEPEPTFAMELDDLPKRLKELIFOE 323
Db 382 PYLEQYDPGDEPVCCEPFTLEMEFDDLPKRLKELIWE 421

RESULT 15
T51944
pathogenicity MAP kinase 1 [imported] - Pyricularia grisea
C:Species: Pyricularia grisea
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T51944
R:Xu, J.R.; Hamer, J.E.
Genes Dev. 10, 2696-2706, 1996
A:Title: MAP kinase and CAMP signaling regulate infection structure formation and pathog
A:Reference number: Z25880; MUID:97102671; PMID:8946911
A:Accession: T51944
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-356 <XU>
A:Cross-references: UNIPROT:Q92246; UNIPARC:UPI000006C824; EMBL:U70134; PIDN:AAC49521.2
A:Experimental source: strain Guy11
C:Genetics:
A:Gene: PMK1
A:Introns: 39/2; 232/1
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 51.4%; Score 907; DB 2; Length 356;
Best Local Similarity 52.2%; Pred. No. 4.1e-34;
Matches 177; Conservative 47; Mismatches 67; Indels 48; Gaps 2;

Qy 36 FDVGPRTYQLOYIGEGAYGVSSAYDHYVTRVAKKISPFPHQTYCQRTLREIQLRF 95
Db 18 FNVSEQYDQDVVGEAGYGVCSAIHKPSGQKVAIKKIPFDHSMCLRLREKWLRYF 77
Qy 96 RHENVIGIRDIRASTLEAMRDVYIVQDLMETDLYKLKSKQSLNDHCYFLYQILRGLK 155
Db 78 NHENIISILDIOKPRSEVFNEVYLQELMETDHRVIRTQDLSDHCQYFIOTLRALK 137
Qy 156 YTHSANVLRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEI 215
Db 138 AMHSANVLRDLKPSNLLNANCDLVCFGLARSAASQENNSGFMTYVATRWYRAPEI 197
Qy 216 MLNSKGYTKSIDWSVGCILAEMLSNRPFPGKHLYDQNLHIL----- 257
Db 198 MLTFKTYTKAIDVMSVGCILAEMLSKGPLFPFGKHLYDQNLHILAVVGSNSNA 257
Qy 258 -----LALDILDRMLTFNPKRITVEEALAHFYLQYD 291
Db 258 RAREYIRSLPFFKKVPFRTLPKTSIDLALDLEKLAFAFPVKITVEEALKHPFLPYHD 317

Qy 292 PTDEPVA---BEPTTFAMELDDLPKRLKELIFQETAR 326
| | | | | | | | : | : | : | : |
Db 318 PDDEFTAPPIPEFFDFDKHKNLSKEQLKQFIQEIMR 356

Search completed: February 6, 2006, 15:35:33
Job time : 26.1009 secs

QY	1	MAAAAAQGGGGGPRTEGVGPGVEVMVWQGPFDVGPYTOLOYTGEGAYGMVSSAY	60
Db	1	MAAAAAQGGGGGPRTEGVGPGVEVMVWQGPFDVGPYTOLOYTGEGAYGMVSSAY	60
QY	61	DHVKTRVAIKLSPSEHOTYQORTLEIQIKLLRFRHENVIGIRDIRASTLEAMRDVYI	120
Db	61	DHVKTRVAIKLSPSEHOTYQORTLEIQIKLLRFRHENVIGIRDIRASTLEAMRDVYI	120
QY	121	VQDLMETDLYKLKLSQQLSNDHI CYFLYQILRGLKYIHSANYLHRDLKPSNLLINTTCOL	180

Db 121 VQDLMETDYLKLLKSQSLNDHICYFLYQILRGLKVHISANVLHRLDKPSNLLSNTTCDL 180
 Qy 181 KICDFGLARIADPHDHTGFLTYVATRWYRAPEIMLNSKGTYKSDIWSVGCILAEML 240
 Db 181 KICDFGLARIADPHDHTGFLTYVATRWYRAPEIMLNSKGTYKSDIWSVGCILAEML 240
 Qy 241 NRPIFGKHVLDQNLHILALDLRLMLTFNPNKRITVEEALAHPLYLEQYDPTDEPVAEE 300
 Db 241 NRPIFGKHVLDQNLHILALDLRLMLTFNPNKRITVEEALAHPLYLEQYDPTDEPVAEE 300
 Qy 301 PPTFAMELDDLPKRLKELIFQSTARFQGVLEAP 335
 Db 301 PPTFAMELDDLPKRLKELIFQSTARFQGVLEAP 335
 RESULT 2
 MK03_HUMAN STANDARD; PRT; 379 AA.
 AC PW7361;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
 DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
 DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
 DE associated protein 2 kinase)
 DE Names=MAPK3; Synonyms=ERK1, PRK3;
 GN Homo sapiens (Human).
 OS
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Hepatoma;
 RX MEDLINE=93330262; PubMed=7687743;
 RA Charest D.L., Jirik P., Harder K., Pelech S.L., Mordret G.;
 RT "Molecular cloning, expression, and characterization of the human
 RT mitogen-activated protein kinase p44erk1.";
 RL Mol. Cell. Biol. 13:4679-4690(1993).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighan S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 14-379.
 RX MEDLINE=92171961; PubMed=1540184;
 RA Owaki H., Makar R., Boulton T.G., Cobb M.H., Geppert T.D.;
 RT "Extracellular signal-regulated kinases in T cells: characterization
 RT of human ERK1 and ERK2 cDNAs.";
 RL Biochem. Biophys. Res. Commun. 182:1416-1422(1992).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 25-379.

RX MEDLINE=92316223; PubMed=1319925; DOI=10.1016/0014-5793(92)80612-K;
 RA Gonzalez F.A., Raden D.L., Rigby M.R., Davis R.J.;
 RT "Heterogeneous expression of four MAP kinase isoforms in human
 RT tissues.";
 RL FEBS Lett. 304:170-178(1992).
 RN [5]
 RP INTERACTION WITH HIV-1 NEF.
 RX PubMed=8794306;
 RA Greenway A.L., Azad A., Mills J., McPhee D.A.;
 RT "Human immunodeficiency virus type 1 Nef binds directly to LCK and
 RT mitogen-activated protein kinase, inhibiting kinase activity.";
 RL J. Virol. 70:6701-6708(1996).
 CC !- FUNCTION: Involved in both the initiation and regulation of
 CC meiosis, mitosis, and postmitotic functions in differentiated
 CC cells by phosphorylating a number of transcription factors such as
 CC Elk-1. Phosphorylates EIF4EBP1; required for initiation of
 CC translation. Phosphorylates microtubule-associated protein 2
 CC (MAP2) (By similarity).
 CC !- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC !- COFACTOR: Magnesium (By similarity).
 CC !- ENZYME REGULATION: Activated and tyrosine phosphorylated in
 CC response to insulin and NGF. This interaction inhibits its kinase
 CC !- SUBUNIT: Binds to HIV-1 Nef. This interaction inhibits its kinase
 CC activity.
 CC !- INTERACTION:
 CC P53355:DAPK1; NbExp=3; IntAct=EBI-73995, EBI-358616;
 CC O75676:RPS6KA4; NbExp=1; IntAct=EBI-73995, EBI-73933;
 CC !- PTM: Autophosphorylated on threonine and tyrosine residues (By
 CC similarity).
 CC !- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
 CC kinase subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; X60188; CAA42744.1; -; mRNA.
 CC DR EMBL; BC013992; AAH13992.1; -; mRNA.
 CC DR EMBL; M84490; AAA36142.1; -; mRNA.
 CC DR EMBL; Z11696; CAA77754.1; -; mRNA.
 CC DR PIR; A48082; A48082.
 CC DR HSSP; P28482; 1PME.
 CC DR SMR; P27361; 29-374.
 CC DR InAct; P27361; -.
 CC DR Ensembl; ENSG00000102892; Homo sapiens.
 CC DR HGNC; HGNC:16877; MAPK3.
 CC DR H-InvDB; HIX0012930; -.
 CC DR Reactome; P27361; -.
 CC DR MIM; 601795; -.
 CC DR GO; GO:0005524; F-ATP binding; NAS.
 CC DR GO; GO:0004707; F-MAP kinase activity; NAS.
 CC DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 CC DR GO; GO:0000074; P:regulation of cell cycle; NAS.
 CC DR InterPro; IPR008349; Erk 1-2 MAPK.
 CC DR InterPro; IPR008351; JNK_MAPK.
 CC DR InterPro; IPR003527; MAP_kin.
 CC DR InterPro; IPR000719; Prot_kinase.
 CC DR InterPro; IPR008271; Ser_thr_pkin_AS.
 CC DR InterPro; IPR002290; Ser_thr_pkinase.
 CC DR Pfam; PF00069; Pkinase; 1.
 CC DR PRINTS; PR01770; ERK1ERK2MAPK.
 CC DR PRINTS; PR01772; JNKMAPKINASE.
 CC DR ProDom; PD000001; Prot_kinase; 1.
 CC DR SMART; SM00220; S_TK; 1.
 CC DR PROSITE; PS01351; MAPK; 1.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 42 330 Protein kinase.

FT NP_BIND 48 56 ATP (By similarity).
 FT ACT_SITE 166 166 Proton acceptor (By similarity).
 FT BINDING 71 71 ATP (By similarity).
 FT MOD_RES 202 202 Phosphothreonine (activates the kinase).
 FT MOD_RES 204 204 Phosphotyrosine (activates the kinase).
 FT CONFLICT 174 174 I -> S (in Ref. 1).
 SQ SEQUENCE 379 AA; B6020CE413EC41F7 CRC64;
 Query Match 98.2%; Score 1731; DB 1; Length 379;
 Best Local Similarity 88.4%; Pred. No. 1-7e-100;
 Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAAAAAGGGGPRTEGVGPGEVEMVKQDPDPRYTQLOVIGEGYGMVSSAY 60
 DB 1 MAAAAAGGGGPRTEGVGPGEVEMVKQDPDPRYTQLOVIGEGYGMVSSAY 60
 QY 61 DHVYKTRVAIKKISPEHQYTCORTREIQILFRHENVIGIRDIRASTLEAMRDVYI 120
 DB 61 DHVYKTRVAIKKISPEHQYTCORTREIQILFRHENVIGIRDIRASTLEAMRDVYI 120
 QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGLKTHSANVLRDLKPSNLLINTTCDL 180
 DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGLKTHSANVLRDLKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDTGFLTEYVATRWTRAPIMLSNGYTKSIDINSVGCILAEMLS 240
 DB 181 KICDFGLARIADPEHDTGFLTEYVATRWTRAPIMLSNGYTKSIDINSVGCILAEMLS 240
 QY 241 NRPIFGKHLYDQNLHIL----- 258
 DB 241 NRPIFGKHLYDQNLHIL----- 258
 QY 259 --ALDLDRLMTFNPKRIITVEALAHPLYEQYDPTDEPVAEPEPTFAMELDDLPKRL 316
 DB 301 SKALDLDRLMTFNPKRIITVEALAHPLYEQYDPTDEPVAEPEPTFAMELDDLPKRL 360
 QY 317 KELIFQETARFQGVLEAP 335
 DB 361 KELIFQETARFQGVLEAP 379

RESULT 3
 MK03 MOUSE STANDARD; PRT; 380 AA.
 ID_Q63844; Q61531; ORK0X5; Q91WY5;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 25-OCT-2004 (Rel. 45, last sequence update)
 DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
 DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
 DE (MAP kinase 1) (MAPK1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
 DE associated protein 2 kinase) (MNK1).
 GN Name=Mapk3; Synonyms=Erk1, Pmk3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=FVB/N; TISSUE=Kidney, and Mammary tumor;
 RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Roha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalak U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 4-380.
 RX PubMed=8424957; DOI=10.1016/0167-4781(93)90074-N;
 RA Tanner B., Mueckler M.;
 RT "Molecular cloning of a mouse extracellular signal regulated kinase
 RT (erk-1).";
 RL Biochim. Biophys. Acta 1171:319-320 (1993).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 10-19.
 RC TISSUE=Pre-B cell;
 RX MEDLINE=92020947; PubMed=1717989;
 RA Crews C.M., Alessandrini A.A., Erikson R.L.;
 RT "Mouse Erk-1 gene product is a serine/threonine protein kinase that
 RT has the potential to phosphorylate tyrosine.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8845-8849 (1991).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 77-380.
 RC TISSUE=Fetal brain;
 RX MEDLINE=91369479; PubMed=1716439;
 RA de Miguel C., Kligman D., Patel J., Detexa-Wadleigh S.D.;
 RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
 RT from mouse and rat brain.";
 RL DNA Cell Biol. 10:505-514 (1991).
 RN [5]
 RP NUCLEOTIDE SEQUENCE OF 171-209.
 RC STRAIN=CBA; TISSUE=Bone marrow;
 RX MEDLINE=93185941; PubMed=8444355; DOI=10.1016/0378-1119(93)90411-U;
 RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
 RT "Novel CDC2-related protein kinases produced in murine hematopoietic
 RT stem cells.";
 RL Gene 124:305-306 (1993).
 RN [6]
 RP NUCLEOTIDE SEQUENCE OF 171-205.
 RX MEDLINE=93092802; PubMed=1459009;
 RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
 RT "Identification of new protein kinase genes, similar to kinases of the
 RT cdc2 family and expressed in murine hematopoietic stem cells.";
 RL Dokl. Akad. Nauk SSSR 324:893-897 (1992).
 CC -1- FUNCTION: Involved in both the initiation and regulation of
 CC meiosis, mitosis, and postmitotic functions in differentiated
 CC cells by phosphorylating a number of transcription factors such as
 CC ELK-1. Phosphorylates E1F4EBP1; required for initiation of
 CC translation. Phosphorylates microtubule-associated protein 2
 CC (MAP2) (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium (By similarity).
 CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
 CC phosphorylation.
 CC -1- PTM: Autophosphorylated on threonine and tyrosine residues.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
 CC kinase subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; BC013754; AAH13754.1; -; mRNA.
 CC EMBL; BC029712; AAH29712.1; -; mRNA.
 CC EMBL; S58470; AAB19973.1; -; mRNA.
 CC EMBL; X64605; CAA45889.1; -; mRNA.
 CC PIR; S28184; S28184.
 CC HSSP; P28482; 1PME.

FT	ACT_SITE	167	167	Proton acceptor (By similarity).
FT	BINDING	72	72	ATP (By similarity).
FT	MOD_RES	203	203	Phosphothreonine (activates the kinase)
FT				(By similarity).
FT	MOD_RES	205	205	Phosphotyrosine (activates the kinase)
FT				(By similarity).
FT	VARSPIC	340	340	E -> EVSRPPAAGRGISVPSVRVPVCLCPQ (in isoform B).
FT				/FTid=VSP_004830.
FT	CONFLICT	95	95	G -> R (in Ref. 4).
FT	SEQUENCE	380 AA;	42981 MW;	53785C4ED72723A CRC64;
Query Match				
Best Local Similarity				
Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;				
Qy	2	AAAAAGGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTYQLQYIGEGAYGMVSSAYD	61	
Db	3	AAAAAPGGGGGPRGTAGVVPVGPGEVEMVKGPFDVGPRTYQLQYIGEGAYGMVSSAYD	62	
Qy	62	HVKTRVAIKKISPFHEQTYCQRTLRBIQILLRPHENVIGIRDLRPASTLEAMRDYIV	121	
Db	63	HVKTRVAIKKISPFHEQTYCQRTLRBIQILLRPHENVIGIRDLRPASTLEAMRDYIV	122	
Qy	122	QDLMETDLYKLLKSKQSLNDHICFVLYQIILRGUKYIHSANVLRDLKPSNLLINTTCDLK	181	
Db	123	QDLMETDLYKLLKSKQSLNDHICFVLYQIILRGUKYIHSANVLRDLKPSNLLINTTCDLK	182	
Qy	182	ICDFGLARIADPEHDHTGFLTEVATWYRAPEIMLSKGYTKSIDIWSVGCILAELMSN	241	
Db	183	ICDFGLARIADPEHDHTGFLTEVATWYRAPEIMLSKGYTKSIDIWSVGCILAELMSN	242	
Qy	242	RP1PPGKHYLDQLNHIL-----	258	
Db	243	RP1PPGKHYLDQLNHILGLSPSQEDLNCINMKARNYLQSLPSKTKVAKLFPKSDS	302	
Qy	259	-ALDILDRMLTFPNPKKITVEEALAHPLYEQYDPTDEPVAEEBFTFAMELDDLKPERLK	317	
Db	303	KALDILLDRMLTFPNPKKITVEEALAHPLYEQYDPTDEPVAEEBFTFAMELDDLKPERLK	362	
Qy	318	ELIFQETARFQPGVLEAP 335		
Db	363	ELIFQETARFQPGVLEAP 380		
RESULT 5				
Q4PIY8_RAT				
ID	Q4PIY8_RAT	PRELIMINARY;	PRT;	380 AA.
AC	Q4PIY8;			
DT	13-SEP-2005	(TrEMBLrel. 31, Created)		
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)		
DE	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)		
DE	Extracellular-signal-regulated kinase 1.			
GN	Name=Mapk3;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=93013050; PubMed=1327976; DOI=10.1016/0378-1119(92)90109-3;			
RA	Marquardt B., Stabel S.;			
RT	"Sequence of a rat cDNA encoding the ERK1-MAP kinase.";			
RL	Gene 120:297-299(1992).			
DR	EMBL; S46779; AA011604.1; -; mRNA.			
KW	Kinase.			
SQ	SEQUENCE	380 AA;	42981 MW;	53785C4ED72723A CRC64;
Query Match				
Best Local Similarity				
Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;				

QY 2 AAAAAAGGGGEGRRTEGVGPGVGEVEMVKGQFDVGPRTYQLOVIGEGAYGMWSSAYD 61
Db 3 AAAAAAGGGGEGRRTAGVVPVPGVEVEMVKGQFDVGPRTYQLOVIGEGAYGMWSSAYD 62
QY 62 HVKTRVAIKKISPFHQYTCQRTLEIQILRRFHHENVIGIRDLIRASTLEAMRDVYIV 121
Db 63 HVKTRVAIKKISPFHQYTCQRTLEIQILRRFHHENVIGIRDLIRASTLEAMRDVYIV 122
QY 122 QDLMETDLYKLLKSQQLSNDHICVFLYQILRGKLYIHSANVLHRDLKPSNLLINTTCDLK 181
Db 123 QDLMETDLYKLLKSQQLSNDHICVFLYQILRGKLYIHSANVLHRDLKPSNLLINTTCDLK 182
QY 182 ICDPGLARTADPHDHTGFLTEVATRWYRAPIMLSNGYTKSIDIWSVGCILAEMLSN 241
Db 183 ICDPGLARTADPHDHTGFLTEVATRWYRAPIMLSNGYTKSIDIWSVGCILAEMLSN 242
QY 242 RPIFPKGHYLDQNLHIL----- 258
Db 243 RPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPKTKVAAKLPKSDS 302
QY 259 -ALDLDRLMTFNPKNRITVEEALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLK 317
Db 303 KALDLDRLMTFNPKNRITVEEALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLK 362
QY 318 ELIFQETARFQGVLEAP 335
Db 363 ELIFQETARFQGAPEAP 380

RESULT 6

ID Q7M0H9 CRIGR PRELIMINARY; PRT; 369 AA.
AC Q7M0H9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) ERK1 (Fragment).
DE Cricetulus griseus (Chinese hamster).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Cricetidae; Cricetinae; Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92199340; PubMed=1372523;
RA Meloche S., Pages G., Pouyssegur J.;
RT "Functional expression and growth factor activation of an epitope-tagged p44 mitogen-activated protein kinase, p44(mak).";
RL Mol. Biol. Cell 3:63-71(1992).
CC -1- FUNCTION: Kinase involved in a signal transduction pathway that is activated by changes in the osmolarity of the extracellular environment. Plays an essential role in maintaining water homeostasis (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC PIR; A60041; A60041.
DR HSSP; Q16539; IOZI.
DR SMR; Q7M0H9; 19-364.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k. . .; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP kin.
DR InterPro; IPR00719; Prof kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; I.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase; 1.

DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-Binding; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1 369
SQ SEQUENCE 369 AA; 42126 MW; 0BB8FCAAC690B4FF CRC64;

Query Match 92.8%; Score 1636; DB 2; Length 369;
Best Local Similarity 85.6%; Pred. No. 1.4e-94;
Matches 316; Conservative 2; Mismatches 7; Indels 44; Gaps 1;

QY 11 GSEPRTEGVGPGVGEVEMVKGQFDVGPRTYQLOVIGEGAYGMWSSAYDHVRKTRVAI 70
Db 1 GGEPRGAAGVPGVGEVEMVKGQFDVGPRTYQLOVIGEGAYGMWSSAYDHVRKTRVAI 60
QY 71 KKISPFHQYTCQRTLEIQILRRFHHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLY 130
Db 61 KKISPFHQYTCQRTLEIQILRRFHHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLY 120
QY 131 KLLKSQQLSNDHICVFLYQILRGKLYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARI 190
Db 121 KLLKSQQLSNDHICVFLYQILRGKLYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARI 180
QY 191 ADPEHDHTGFLTEVATRWYRAPIMLSNGYTKSIDIWSVGCILAEMLSNRPFPKGHY 250
Db 181 ADPEHDHTGFLTEVATRWYRAPIMLSNGYTKSIDIWSVGCILAEMLSNRPFPKGHY 240
QY 251 LDQLNLHIL-----ALDLDRLM 266
Db 241 LDQLNLHILGILGSPSQEDLNCIINMKARNYLQSLPKTKVAAKLPKSDSKALDLDRLM 300
QY 267 LTFNPNKRITVEEALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLKELIQETAR 326
Db 301 LTFNPNKRITVRDALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLKELIQETAR 360
QY 327 FQGVLEAP 335
Db 361 FQGAPEAP 369

RESULT 7
Q8NHX1 HUMAN
ID Q8NHX1_HUMAN PRELIMINARY; PRT; 357 AA.
AC Q8NHX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-related kinase 1b.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Aebersold D.M., Yung Y., Seger R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033607; AAK52329.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q8NHX1; 29-344.
DR Ensemble; ENSG00000102882; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP kin.
DR InterPro; IPR000719; Prot_kinase.
DR ProDom; PD000001; Prot_kinase; 1.

DR	Pfam; PF00069; Pkinase; 1.	DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR	PRINTS; PR01770; ERK1ERK2MAPK.	DR	InterPro; IPR008349; Erk_1_2_MAPK.
DR	PRINTS; PR01772; JNKMAPKKNASE.	DR	InterPro; IPR003527; MAP_kin.
DR	ProDom; PD000001; Prot_kinase; 1.	DR	InterPro; IPR00719; Prot_kinase.
DR	SMART; SM00220; S_TKC; 1.	DR	InterPro; IPR008271; Ser_thr_pkin_AS.
DR	PROSITE; PS01351; MAPK; UNKNOWN 1.	DR	Pfam; PF00069; Pkinase; 1.
DR	PROSITE; PS0107; PROTEIN KINASE ATP; 1.	DR	PRINTS; PR01770; ERK1ERK2MAPK; 2.
DR	PROSITE; PS00107; PROTEIN KINASE DOM; 1.	DR	ProDom; PD000001; Prot_kinase; 2.
DR	PROSITE; PS0011; PROTEIN KINASE ST; 1.	DR	PROSITE; PS01351; MAPK; 1.
KW	ATP-binding; Kinase; Nucleotide-Binding;	DR	PROSITE; PS00108; PROTEIN KINASE-ST; 1.
KW	Serine/threonine-protein kinase; Transferase.	KW	ATP-binding; kinase; Nucleotide-Binding;
SQ	SEQUENCE 357 AA; 40062 MW; 58C92773988ADA79 CRC64;	KW	Serine/chronine-protein kinase; transferase.
Query Match 86.4%; Score 1522.5; DB 2; Length 357;			
Best Local Similarity 85.8%; Pred. No. 1.8e-87;			
Matches 296; Conservative 1; Mismatches 3; Indels 45; Gaps 2;			
QY	1 MAAAAQGGGGGPRRTGEGVGPGEVEMVKQPPFDVGPRTYQYIGEGAYGMVSSAY 60	QY	3 AAAAAQGGGGGPRRTGEGVGPGEVEMVKQPPFDVGPRTYQYIGEGAYGMVSSAYDH 62
DB	1 MAAAAAQQGGGGGPRRTGEGVGPGEVEMVKQPPFDVGPRTYQYIGEGAYGMVSSAY 60	DB	25 AAVAVGGANG-----AAGP-KPG-LESVKQGNFDVGPRTYQYIGEGAYGMVCSAFDN 76
QY	61 DHVRKTRVAIKKISPPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	QY	63 VNRKTRVAIKKISPPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 122
DB	61 DHVRKTRVAIKKISPPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	DB	77 VNRKTRVAIKKISPPFHQYTCORTLREIKILLRFRHENVIGIRDILRASTLEAMRDVYI 136
QY	121 VDLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180	QY	123 DLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI 182
DB	121 VDLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180	DB	137 DLMETDLYKLLKIQQLSNDHVCYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI 196
QY	181 KICDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240	QY	183 CDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLSNR 242
DB	181 KICDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240	DB	197 CDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLSNR 256
QY	241 NRPTFGKHLYDQLNHIL----- 258	QY	243 PIFPGKHLYDQLNHIL----- 258
DB	241 NRPTFGKHLYDQLNHILGILGSPSQEDLNCIINMKARNYQLSPKTKVAKLPPKSD 300	DB	257 PIFPGKHLYDQLNHILGILGSPSQEDLNCIINMKARNYQLSPKTKIPWPKLFPKADNK 316
QY	259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPP 301	QY	259 ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERLKE 318
DB	301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDE-VGQSP 344	DB	317 ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERLKE 376
RESULT 8			
ID	Q7LZH4_CYPCHA	ID	Q8UMG6_CHICK
AC	Q7LZH4_CYPCHA PRELIMINARY; PRT; 392 AA.	AC	Q8UMG6_CHICK PRELIMINARY; PRT; 368 AA.
DT	01-MAR-2004 (TrEMBLrel. 26, Created)	DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)	DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-FEB-2005 (TrEMBLrel. 29, Last annotation update)	DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Extracellular signal-regulated kinase (EC 2.7.-.-) 1 (Extracellular	DE	Extracellular signal-regulated kinase 2.
DE	signal regulated protein kinase 1).	OS	Gallus gallus (Chicken).
GN	Name=cERK1;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OS	Cyprinus carpio (Common carp).	OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Gallus.
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	OC	Gallus.
OC	Cyprinidae; Cyprinus.	OX	NCBI_TaxID=9031;
OX	NCBI_TaxID=7962;	RN	[1]
RN	[1]	RP	NUCLEOTIDE SEQUENCE.
RP	NUCLEOTIDE SEQUENCE.	RC	TISSUE=Pinel gland.
RC	TISSUE=Ovary;	RA	Kasahara T., Higashi C., Okano T., Fukada Y.;
RX	MEDLINE=98269030; PubMed=9603989;	RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RA	Hashimoto H., Yokoyama Y., Matsuo Y., Toyohara H., Kohno M.,	CC	- - CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
RA	Sakaguchi M.;	CC	- - SIMILARITY: Belongs to the Ser/Thr protein kinase family.
RT	"Existence of two isoforms of extracellular signal-regulated kinase in	DR	EMBL; AY013635; AAK56503.1; -; mRNA.
RT	fish.";	DR	HSSP; Q16539; 1021.
RL	J. Biochem. 123:1031-1035(1998).	DR	SMR; Q7LZH4; 39-386.
CC	- - SIMILARITY: Belongs to the Ser/Thr protein kinase family.	DR	GO; GO:0005524; F:ATP binding; IEA.
DR	EMBL; AB006038; BAD23842.1; -; mRNA.	DR	GO; GO:0004707; F:MAP kinase activity; IEA.
DR	PIR; J00052; J00052.	DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR	HSSP; Q16539; 1021.		
DR	SMR; Q7LZH4; 39-386.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004707; F:MAP kinase activity; IEA.		
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.		

QY	3	AAAAGGGGEGRRRTTEGVGPGVEMVKGPFDVGPRTYTOLOQYIGEGAGVMSAYDH	62
DB	18	SSAAGPGGAVAPGGPSGAAGSKFG-LESVKQGNFVGPRTYDLOQYIGEGAGVMSAFDN	76
QY	63	VRKTRVAIKKISPFHEQTYCQRTLEIRIQIILLRPHENVIGIRDIRLRASTLEAMRDYVIVQ	122
DB	77	VNKIRVAIKKISPFHEQTYCQRTLEIRIKILLRPHENIIGINDILRARHIDYMRDVIYVQ	136
QY	123	DLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDLKI	182
DB	137	DLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDLKI	196
QY	183	CDLGLARIADPEHDHTGFLTEYVATWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR	242
DB	197	CDLGLARIADPEHDHTGFLTEYVATWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR	256
QY	243	PIPPGKHLYDLQNLNHL-----	258
DB	257	PIPPGKHLYDLQNLNHL-----	316
QY	259	ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVAEBPFTFAMELDDLKPKERLKE	318
DB	317	ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVAEBPFTFAMELDDLKPKERLKE	376
QY	319	LIFOETARFQ 328	
DB	377	LIFEETARFQ 386	

QY	3	AAAAGGGGEGRRRTTEGVGPGVEMVKGPFDVGPRTYTOLOQYIGEGAGVMSAYDH	62
DB	17	SSAAGPGGAVAPGGPSGAAGSKFG-LESVKQGNFVGPRTYDLOQYIGEGAGVMSAFDN	75
QY	63	VRKTRVAIKKISPFHEQTYCQRTLEIRIQIILLRPHENVIGIRDIRLRASTLEAMRDYVIVQ	122
DB	76	VNKIRVAIKKISPFHEQTYCQRTLEIRIKILLRPHENIIGINDILRARHIDYMRDVIYVQ	135
QY	123	DLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDLKI	182
DB	136	DLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDLKI	195
QY	183	CDLGLARIADPEHDHTGFLTEYVATWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR	242
DB	196	CDLGLARIADPEHDHTGFLTEYVATWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR	255
QY	243	PIPPGKHLYDLQNLNHL-----	258
DB	256	PIPPGKHLYDLQNLNHL-----	315
QY	259	ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVAEBPFTFAMELDDLKPKERLKE	318
DB	316	ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVAEBPFTFAMELDDLKPKERLKE	375
QY	319	LIFOETARFQ 328	
DB	376	LIFEETARFQ 385	

Query Match 82.1%; Score 1447.5; DB 2; Length 391;
 Best Local Similarity 76.5%; Pred. No. 9.6e-83;
 Matches 283; Conservative 14; Mismatches 28; Indels 45; Gaps 2;

RESULT 12

MK01_BOVIN

ID AC P46196; STANDARD; PRT; 360 AA.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE 13-SEP-2005 (Rel. 48, Last annotation update)

DE Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular

DE signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase

DE 2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERK1).

GN Name=MAPK1; Synonyms=ERK2, PRKMI;

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

OC Pecora; Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Adrenal medulla;

RA Ely C.M., Cox M.E., Her J., Parsons S.J.;

RT "Cloning and sequencing of ERK2 from a bovine adrenal medulla cDNA

RT library.";

RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Phosphorylates microtubule-associated protein 2 (MAP2).

CC Myelin basic protein (MBP), and Elk-1; may promote entry in the

CC cell cycle.

CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.

CC -1- COFACTOR: Magnesium (by similarity).

CC -1- ENZYME REGULATION: Activated by phosphorylation on tyrosine and

CC threonine in response to insulin and NGF.

CC -1- PTM: Autophosphorylated on threonine and tyrosine residues, which

CC correlates with a slow and low level of autoactivation (By

CC similarity).

CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP

CC kinase subfamily.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

RESULT 11

Q9DGR6_BRARE

ID Q9DGR6 BRARE PRELIMINARY; PRT; 391 AA.

AC Q9DGR6; (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ERK1.

GN Name=mapk3; Synonyms=ZERK1;

OS Brachydanio rerio (Zebrafish)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hirano T., Fujii R., Hibi M.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030902; BAB11812.1; -; mRNA.

DR HSP; P28482; 1PME.

DR SMR; Q9DGR6; 40-385.

DR ZFIN; ZDR-GENE-040121-1; mapk3.

DR GO; GO:000524; F:ATP binding; IEA.

DR GO; GO:0004707; F:MAP kinase activity; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR008349; Erk 1.2 MAPK.

DR InterPro; IPR008351; JNK MAPK.

DR InterPro; IPR003527; MAP_kin.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR008271; Ser_Thr_pkin_AS.

DR InterPro; IPR002290; Ser_Thr_pkinase.

DR Pfam; PF00069; Pkinase; 1.

DR PRINTS; PR01770; ERK1ERK2MAPK.

DR PRINTS; PR01772; JNKMAPKINASE.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKc; 1.

DR PROSITE; PS01351; MAPK; 1.

DR PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

SQ SEQUENCE 391 AA; 44126 MW; 6638F0807B5F9889 CRC64;

[illegible]

RC STRAIN=Swiss; TISSUE=Fibroblast;
RX MEDLINE=91305126; PubMed=1649458;
RA Her J.-H., Wu J.-S., Kall T.B., Sturgill T.W., Weber M.J.;
RT "Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by
RL tyrosine phosphorylation.";
RN Nucleic Acids Res. 19:3743-3743 (1991).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J, and NOD; TISSUE=Head, Thymus, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RX Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bratt D., Brusic V., Clothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Glessi C., Godzik A., Gough J.,
RA Grimmond S., Gustinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongsaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltails L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN Nature 420:563-573 (2002).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Eye;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[4]
RP NUCLEOTIDE SEQUENCE OF 151-189.
RC STRAIN=CBA; TISSUE=Bone marrow;
RX MEDLINE=93185941; PubMed=8444355; DOI=10.1016/0378-1119(93)90411-U;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Novel CDC2-related protein kinases produced in murine hematopoietic
RT stem cells.";
RL Gene 124:305-306 (1993).
[5]
RP PHOSPHORYLATION SITES THR-183 AND TYR-185, AND PARTIAL PROTEIN

RP SEQUENCE.
RX MEDLINE=91184134; PubMed=1849075;
RA Payne D.M., Rossomando A.J., Martino P., Erickson A.K., Her J.-H.,
RA Shabanowitz J., Hunt D.F., Weber M.J., Sturgill T.W.;
RT "Identification of the regulatory phosphorylation sites in
RT pp42/mitogen-activated protein kinase (MAP kinase).";
RL EMBO J. 10:885-892 (1991).
CC -!- FUNCTION: Involved in both the initiation and regulation of
CC meiosis, mitosis, and postmitotic functions in differentiated
CC cells by phosphorylating a number of transcription factors such as
CC ELK-1. Phosphorylates E1F4BPI; required for initiation of
CC translation. Phosphorylates microtubule-associated protein 2
CC (MAP2) (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
CC threonine in response to insulin and NGF. Both phosphorylations
CC are required for activity (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- PTM: Autophosphorylated on threonine and tyrosine residues, which
CC correlates with a slow and low level of autoactivation (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL: X58712; CAA41548.1; -; mRNA.
CC EMBL: AK035386; BAC29053.1; -; mRNA.
CC EMBL: AK048127; BAC33251.1; -; mRNA.
CC EMBL: AK087925; BAC40044.1; -; mRNA.
CC EMBL: BC058258; AAH58258.1; -; mRNA.
CC EMBL: D10939; BAA01733.1; -; mRNA.
CC PIR: S16444; S16444.
CC SMR: P63085; 14-357.
CC IntAct: P63085; -.
CC MGI: MGI:1346858; Mapk1.
CC GO: GO:0005737; C:cytoplasm; IDA.
CC GO: GO:0005634; C:nucleus; IDA.
CC GO: GO:0004707; F:MAP kinase activity; IDA.
CC GO: GO:0001784; F:phosphotyrosine binding; IMP.
CC GO: GO:0005515; F:protein binding; IPI.
CC GO: GO:0004672; F:protein kinase activity; IDA.
CC GO: GO:0000165; P:MAPKK cascade; IDA.
CC GO: GO:0009887; P:organogenesis; IDA.
CC GO: GO:0006468; P:protein amino acid phosphorylation; IDA.
CC GO: GO:0006974; P:response to DNA damage stimulus; IDA.
CC GO: GO:0007165; P:signal transduction; TAS.
CC InterPro: IPR008349; Erk_1_2_MAPK.
CC InterPro: IPR003527; MAP_kin.
CC InterPro: IPR008271; Ser_Thr_pkin_AS.
CC Pfam: PF00069; Pkinase; I.
CC PRINTS: PR01770; ERKLERK2MAPK.
CC PROSITE: PS01351; MAPK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00109; Cell cycle: Direct protein sequencing; Kinase;
CC Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase;
CC Transferase.
CC DOMAIN 23 311 Protein kinase.
CC NP_BIND 29 37 ATP (By similarity).
CC COMPBIAS 2 37 Poly-Ala.
CC ACT_SITE 147 147 By similarity.
CC BINDING 52 52 ATP (By similarity).
CC MOD_RES 183 183 Phosphothreonine (activates the kinase).
CC MOD_RES 185 185 Phosphotyrosine (activates the kinase).
CC SEQUENCE 358 AA; 41276 MW; 3B8CF22471EDBA0B CRC64;

	Query Match	81.4%; Score 1434.5; DB 1; Length 358;
	Best Local Similarity	73.8%; Pred. No. 5.7e-82;
	Matches 276; Conservative	19; Mismatches 16; Indels 63; Gaps 2;
QY	1	MAAAAAGCGGGGPRRTSGVPGVEVMKGPDPGVRYTOLQVIGGAYGMVSSAY 60
DB	1	MAAAAAG-----PEMVRGQVDVGPRYTSLVSIGGAYGMVCSAY 41
QY	61	DHVRTRVAIKKISPFHQTYCQRTLRIQILLRFRHENVIIRDLIRASTLEAMRDVYI 120
DB	42	DNLNKRVVAIKKISPFHQTYCQRTLREIKILLRFRHENIIGINDIIRAFTIEQMKVYI 101
QY	121	VQDLMETDLYKLLKSQQISNDHI CYFLYQVILRGGLKYTHSANVLHRDLKPSNLLINTTCDL 180
DB	102	VQDLMETDLYKLLKTQHLSNDHI CYFLYQVILRGGLKYTHSANVLHRDLKPSNLLINTTCDL 161
QY	181	KICDFGLARIADPEHDHTGFELTEYVATWYRAPEIMLNSKGYSKIIDSWSGCCILAEMLS 240
DB	162	KICDFGLARVADPDHDHTGFUTEYVATWYRAPEIMLNSKGYSKIIDSWSGCCILAEMLS 221
QY	241	NRPIDFGKHLDQLNHIL----- 258
DB	222	NRPIDFGKHLDQLNHILGILGPSQEDLNCINLKARNYLSSLPHKNKVPWNRLFPENAD 281
QY	259	--ALDLDRMLTFNPKNKITVEEALAHPLYEQYDPTDEPVAEPFFTFAMELDDLPKERL 316
DB	282	SKALOLLDKMLTFNPKRIEVEQALAHPLYEQYDPSDEPTAEAPFKFDMELDDLPERKL 341
QY	317	KELIFQETARFQPG 330
DB	342	KELIFEETARFQPG 355

RESULT 15

ID	MKOI_RAT	STANDARD;	PRT;	358 AA.
AC	P63086;			
DC	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Mitogen-activated protein kinase 1 (SC 2.7.1.37) (Extracellular signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase 2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERT1).			
GN	Name=Mapk1; Synonyms=Erk2, Mapk, Prkm1;			
OS	Rattus norvegicus (rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.			
OX	NCBI_TaxId=10116;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RC	MEDLINE=91235302; PubMed=2032290; DOI=10.1016/0092-8674(91)90098-J;			
RA	Boulton T.G., Nye S.H., Robbins D.J., Ip N.Y., Radziejewska E., Morgenbesser S.D., Depinho R.A., Panayotatos N., Cobb M.H., Yancopoulos G.D.;			
RA	"ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF.";			
RT	Cell 65:663-675(1991).			
RL	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).			
RC	MEDLINE=94150699; PubMed=8107865; DOI=10.1038/367704a0;			
RA	Zhang F., Strand A., Robbins D., Cobb M.H., Goldsmith E.J.;			
RA	"Atomic structure of the MAP kinase ERK2 at 2.3-A resolution.";			
RT	Nature 367:704-710(1994).			
RL	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).			
RC	MEDLINE=97442349; PubMed=9298898; DOI=10.1016/S0092-8674(00)80351-7;			
RA	Canagarajah B.J., Khokhlatchev A., Cobb M.H., Goldsmith E.J.;			
RA	"Activation mechanism of the MAP kinase ERK2 by dual phosphorylation.";			
RT	Cell 90:859-869(1997).			
RL				

[4] AUTOPHOSPHORYLATION.

XP MEDLINE=91296777; PubMed=1712480;

Seger R., Ahn N.G., Boulton T.G., Yancopoulos G.D., Panayotatos N., Radziejewska E., Ericsson L., Bratlien R.L., Cobb M.H., Krebs E.G.;

"Microtubule-associated protein 2 kinases, ERK1 and ERK2, undergo autophosphorylation on both tyrosine and threonine residues: implications for their mechanism of activation.";

Proc. Natl. Acad. Sci. U.S.A. 88:6142-6146(1991).

[5] PHOSPHORYLATION OF EIF4EBP1.

XP MEDLINE=95025978; PubMed=7939721;

Liu T.-A., Kong X., Haystead T.A.J., Pause A., Belsham G.J., Sonenberg N., Lawrence J.C. Jr.;

"PHAS-I as a link between mitogen-activated protein kinase and translation initiation.";

Science 266:653-656(1994).

-!- FUNCTION: Involved in both the initiation and regulation of meiosis, mitosis, and postmitotic functions in differentiated cells by phosphorylating a number of transcription factors such as ELK-1. Phosphorylates EIF4EBP1; required for initiation of translation. Phosphorylates microtubule-associated protein 2 (MAP2).

-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.

-!- COFACTOR: Magnesium (By similarity).

-!- ENZYME REGULATION: Activated by phosphorylation on tyrosine and threonine in response to insulin and NGF. Both phosphorylations are required for activity.

-!- TISSUE SPECIFICITY: Highest levels within the nervous system, expressed in different tissues, mostly in muscle, thymus and heart.

-!- DEVELOPMENTAL STAGE: Increased expression during development.

-!- PTM: Autophosphorylated on threonine and tyrosine residues, which correlates with a slow and low level of autoactivation.

-!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP kinase subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; M64300; AAA41124.1; -- mRNA.

PIR; A40033; A40033.

PDB; 1ERK; X-ray; @=1-358.

PDB; 1GOL; X-ray; @=1-358.

PDB; 2ERK; X-ray; @=1-358.

PDB; 3ERK; X-ray; @=1-358.

PDB; 4ERK; X-ray; @=1-358.

SMR; P63086; 14-357.

Ensembl; ENSRNOG0000001849; Rattus norvegicus.

Reactome; P63086; --

RGD; 70500; Mapk1

InterPro; IPR008349; Erk_1_2_MAPK.

InterPro; IPR003527; MAP_kin.

InterPro; IPR000719; Prot_kinase.

InterPro; IPR008271; Ser_thr_pkin_AS.

Pfam; PF00069; Pkinase; 1.

PRINTS; PR01770; ERK1ERK2MAPK.

PROSITE; PS01351; MAPK; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

3D-structure; ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase; Transferase.

DOMAIN	23	311
NP_BIND	29	37
COMPBIAS	2	7
ACT_SITE	147	147
BINDING	52	52
MOD_RES	183	183

ATP.

Phosphothreonine (activates the kinase) (By similarity).

Phosphotyrosine (activates the kinase) (By similarity).		185	185
FT	MOD_RES	11	12
FT	STRAND	13	14
FT	TURN	15	16
FT	TURN	20	22
FT	TURN	23	24
FT	STRAND	28	31
FT	STRAND	35	42
FT	STRAND	43	46
FT	TURN	57	58
FT	TURN	60	75
FT	HELI	79	80
FT	TURN	81	81
FT	STRAND	84	89
FT	TURN	93	95
FT	TURN	99	104
FT	STRAND	108	109
FT	STRAND	110	114
FT	HELI	115	116
FT	TURN	121	140
FT	TURN	141	142
FT	STRAND	143	144
FT	HELI	150	152
FT	STRAND	153	155
FT	TURN	157	158
FT	STRAND	161	163
FT	TURN	167	168
FT	STRAND	170	171
FT	HELI	174	176
FT	STRAND	178	178
FT	TURN	181	182
FT	HELI	189	191
FT	HELI	194	196
FT	TURN	197	199
FT	STRAND	200	200
FT	TURN	205	205
FT	HELI	206	221
FT	HELI	231	242
FT	HELI	247	252
FT	HELI	256	264
FT	HELI	273	276
FT	TURN	278	279
FT	HELI	282	291
FT	HELI	296	298
FT	HELI	302	306
FT	TURN	307	307
FT	HELI	309	311
FT	TURN	312	314
FT	HELI	317	319
FT	TURN	330	331
FT	HELI	338	348
FT	TURN	349	349
FT	HELI	350	352
SQ	SEQUENCE	358 AA;	41276 MW; 3BBCF22471EDBA0B CRC64;

Query Match 81.4%; Score 1434.5; DB 1; Length 358;
Best Local Similarity 73.8%; Pred. No. 5.7e-82;
Matches 276; Conservative 19; Mismatches 16; Indels 63; Gaps 2;

Qy	1	MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPPDPVTOLOYIGEGAYGMVSSAY	60
Db	1	MAAAAAG-----PENVRQVDFVDPRTNLSYIGEGAYGMVCSAY	41
Qy	61	DHVKTRVAIKKISPEHQYQCORTLREIQILFRHENVIGIRDLRASTLEAMRDVYI	120
Db	42	DNLNKRVAIKKISPEHQYQCORTLREIKILLFRHENIIGINDIIRAPTIEQMKDVI	101
Qy	121	VQDLMETDLYKLIKSQLSDNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	180
Db	102	VQDLMETDLYKLLKTQHLSDNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLNTTCDL	161

Qy	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVCILAEMLS	240
Db	162	KICDFGLARVADPDHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVCILAEMLS	221
Qy	241	NRPIFPKGHYLDQLNHIL-----	258
Db	222	NRPIFPKGHYLDQLNHILGSPSQEDLNCIINIKARNYLLSLPHKNKVPWNRIFPNAD	281
Qy	259	--ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTFAMELDDDLPKERL	316
Db	282	SKALDLLDRMLTFNPHKRIEVEQALAHPLYEQYDPSDEPTAEAFKFDMLDDDLPKERL	341
Qy	317	KELIFQETARFQPG	330
Db	342	KELIFEETARFQPG	355

Search completed: February 6, 2006, 15:34:38
Job time : 164.777 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:34:56 ; Search time 33.7896 Seconds
(without alignments)
819.670 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAQQGGGGPRRTEGV.....LKELIFQETARFQPGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 0%

Listing first 45 summaries

Database :

- Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1731	98.2	379	2	US-09-411-628-14
2	1731	98.2	379	2	US-10-174-794-14
3	1731	98.2	379	2	US-09-538-092-1021
4	1725	97.8	379	2	US-09-025-580-25
5	1725	97.8	379	2	US-09-642-749-25
6	1725	97.8	631	2	US-09-417-197-39
7	1719	97.5	379	2	US-08-622-277A-8
8	1659	94.1	624	2	US-09-417-197-57
9	1607	91.2	367	1	US-08-176-620A-11
10	1606	91.1	355	2	US-08-461-985-11
11	1606	91.1	355	2	US-08-459-953A-8
12	1595	90.5	365	1	US-09-393-212-8
13	1595	90.5	365	1	US-08-176-620A-2
14	1595	90.5	365	1	US-08-463-862-2
15	1595	90.5	365	1	US-08-461-985-2
16	1595	90.5	365	1	US-08-458-887-2
17	1595	90.5	365	2	US-08-932-787B-2
18	1595	90.5	365	2	US-08-932-012C-2
19	1595	90.5	365	2	US-08-888-818C-2
20	1459.5	82.8	380	2	US-09-949-016-9919
21	1454.5	82.5	380	2	US-08-459-953A-9
22	1454.5	82.5	380	2	US-09-393-212-9
23	1444.5	81.9	360	2	US-09-457-040B-3
24	1444.5	81.9	360	2	US-09-538-092-1028
25	1436	81.5	415	1	US-08-176-620A-4
26	1436	81.5	415	1	US-08-463-862-4
27	1436	81.5	415	1	US-08-461-985-4

28	1436	81.5	415	1	US-08-458-887-4	Sequence 4, Appli
29	1436	81.5	415	2	US-08-932-787B-4	Sequence 4, Appli
30	1436	81.5	415	2	US-08-932-012C-4	Sequence 4, Appli
31	1436	81.5	415	2	US-08-888-818C-4	Sequence 4, Appli
32	1435.5	81.4	360	2	US-08-622-277A-14	Sequence 14, Appl
33	1434.5	81.4	358	1	US-08-176-620A-12	Sequence 12, Appl
34	1434.5	81.4	358	1	US-08-461-985-12	Sequence 12, Appl
35	1434.5	81.4	358	2	US-09-457-040B-4	Sequence 4, Appli
36	1434.5	81.4	364	2	US-09-457-040B-34	Sequence 34, Appl
37	1434.5	81.4	604	2	US-09-417-197-59	Sequence 59, Appl
38	1434.5	81.4	605	2	US-09-417-197-41	Sequence 41, Appl
39	1421.5	80.6	365	2	US-09-457-040B-35	Sequence 35, Appl
40	1415.5	80.3	360	2	US-09-025-580-2	Sequence 2, Appli
41	1415.5	80.3	360	2	US-09-642-749-2	Sequence 2, Appli
42	905	51.3	183	1	US-08-176-620A-8	Sequence 8, Appli
43	905	51.3	183	1	US-08-461-985-8	Sequence 8, Appli
44	905	51.3	183	2	US-08-932-787B-16	Sequence 16, Appl
45	905	51.3	183	2	US-08-932-012C-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-411-628-14

; Sequence 14, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-411-628-14

Query Match	98.2%;	Score 1731;	DB 2;	Length 379;
Best Local Similarity	88.4%;	Pred. No. 1.2e-173;		
Matches	335;	Conservative	0;	Mismatches 0; Indels 44; Gaps 1;
Qy	1	MAAAAQQGGGGPRRTEGV	60	MAAAAQQGGGGPRRTEGV
Db	1	MAAAAQQGGGGPRRTEGV	60	MAAAAQQGGGGPRRTEGV
Qy	61	DHVKTRVAIKTSPFEHQY	120	DHVKTRVAIKTSPFEHQY
Db	61	DHVKTRVAIKTSPFEHQY	120	DHVKTRVAIKTSPFEHQY
Qy	121	VQDLMETDLYKLKSQLSNDHIC	180	VQDLMETDLYKLKSQLSNDHIC
Db	121	VQDLMETDLYKLKSQLSNDHIC	180	VQDLMETDLYKLKSQLSNDHIC
Qy	181	KICDFGLARIADPEHDHTG	240	KICDFGLARIADPEHDHTG
Db	181	KICDFGLARIADPEHDHTG	240	KICDFGLARIADPEHDHTG
Qy	241	NRPIFPCKHYLDQNLHIL	258	NRPIFPCKHYLDQNLHIL
Db	241	NRPIFPCKHYLDQNLHIL	300	NRPIFPCKHYLDQNLHIL
Qy	259	--ALDLDLRLMTFNPKNRIT	316	--ALDLDLRLMTFNPKNRIT
Db	301	SKALDLDLRLMTFNPKNRIT	360	SKALDLDLRLMTFNPKNRIT
Qy	317	KELIFQETARFQPGVLEAP	335	KELIFQETARFQPGVLEAP

```
Db 361 KELIFQETARFQGVLEAP 379
|||||
RESULT 2
US-10-174-794-14
; Sequence 14, Application US/10174794
; Patent No. 6664086
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-174-794-14

Query Match 98.2%; Score 1731; DB 2; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.2e-173;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
121 VDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
121 VDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
241 NRPIFGKHYLDQNLNHL----- 258
241 NRPIFGKHYLDQNLNHLILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLPKSD 300
QY 259 --ALDLLDRMLTNPKNRITVBEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 316
Db 301 SKALDLDLRLMTNPKNRITVBEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379

RESULT 4
US-09-025-580-25
; Sequence 25, Application US/09025580
; Patent No. 6162613
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; APPLICANT: Fox, Ted
; APPLICANT: Willson, Keith Phillip
; APPLICANT: Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
```

```
;
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-580-25

Query Match          97.8%; Score 1725; DB 2; Length 379;
Best Local Similarity 88.1%; Pred. No. 5.1e-173;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRTEGGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGVSSAY 60
DB 1 MAAAAAGGGGGPRTEGGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGVSSAY 60
QY 61 DHVKTRVAIKKISPFHEQTYCORTLREIQILFRFHNHENVIGIRDLASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPFHEQTYCORTLREIQILFRFHNHENVIGIRDLASTLEAMRDVYI 120
QY 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFPKHYLDQNLHIL----- 258
DB 241 NRPIFPKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSLPSKTKVAKLPPKSD 300
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKERL 316
DB 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKERL 360
QY 317 KELIFQTARFQGVLEAP 335
DB 361 KELIFQTARFQGVLEAP 379

RESULT 5
US-09-642-749-25
; Sequence 25, Application US/09642749
; Patent No. 6849716
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; Fox, Ted
; Wilson, Keith Phillip
; Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
; US-09-642-749-25
```

```
Query Match          97.8%; Score 1725; DB 2; Length 379;
Best Local Similarity 88.1%; Pred. No. 5.1e-173;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRTEGGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGVSSAY 60
DB 1 MAAAAAGGGGGPRTEGGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGVSSAY 60
QY 61 DHVKTRVAIKKISPFHEQTYCORTLREIQILFRFHNHENVIGIRDLASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPFHEQTYCORTLREIQILFRFHNHENVIGIRDLASTLEAMRDVYI 120
QY 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFPKHYLDQNLHIL----- 258
DB 241 NRPIFPKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSLPSKTKVAKLPPKSD 300
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKERL 316
DB 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPPTFAMELDDLPKERL 360
QY 317 KELIFQTARFQGVLEAP 335
DB 361 KELIFQTARFQGVLEAP 379
```

```
RESULT 6
US-09-417-197-39
; Sequence 39, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole Thastrup, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
; FILE REFERENCE: 3759-0110P
; CURRENT FILING DATE: 1999-10-07
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 39
```

```

; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-09-417-197-39

Query Match          97.8%; Score 1725; DB 2; Length 631;
Best Local Similarity 88.1%; Pred. No. 1.1e-172;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAY 60
Db 253 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAY 312
QY 61 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 313 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 372
QY 121 VQDLMETDLYKLLKSQQLSNDHHCYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 373 VQDLMETDLYKLLKSQQLSNDHHCYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 432
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
Db 433 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 492
QY 241 NRPIPGKHLYDQNLNHL----- 258
Db 493 NRPIPGKHLYDQNLNHLGILGSPSQEDLNCIINKARNYLSQSPKTKVAVAKLFPKSD 552
QY 259 --ALDLDRLMTFFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 316
Db 553 SKALDLDRLMTFFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 612
QY 317 KELIFQETARFQPGVLEAP 335
Db 613 KELIFQETARFQPGVLEAP 631

RESULT 7
US-08-622-277A-8
; Sequence 8, Application US/08622277A
; Patent No. 6001580
; GENERAL INFORMATION:
; APPLICANT: Tani, Akiyoshi
; APPLICANT: Ichimori, Yuzo
; TITLE OF INVENTION: Method For Assaying MAP Kinase
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,277A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 320577-1995
; FILING DATE: 08-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 305456-1995
; FILING DATE: 24-NOV-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: JP 070125-1995
; FILING DATE: 28-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: toohey, kimberlin M.
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 04221.0039-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
US-08-622-277A-8

Query Match          97.5%; Score 1719; DB 2; Length 379;
Best Local Similarity 87.9%; Pred. No. 2.2e-172;
Matches 333; Conservative 0; Mismatches 2; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPPDVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHHCYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHHCYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIPGKHLYDQNLNHL----- 258
Db 241 NRPIPGKHLYDQNLNHLGILGSPSQEDLNCIINKARNYLSQSPKTKVAVAKLFPKSD 300
QY 259 --ALDLDRLMTFFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 316
Db 301 SKALDLDRLMTFFNPNKRITVAALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 360
QY 317 KELIFQETARFQPGVLEAP 335
Db 361 KELIFQETARFQPGVLEAP 379

RESULT 8
US-09-417-197-57
; Sequence 57, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
US-09-417-197-57

Query Match          94.1%; Score 1659; DB 2; Length 624;

```

```

; TOPOLOGY: unknown
; .MOLSCULE TYPE: protein
US-08-176-620A-11
Query Match 91.2%; Score 1607; DB 1; Length 367;
Best Local Similarity 85.3%; Pred. No. 1.3e-160;
Matches 313; Conservative 1; Mismatches 9; Indels 44; Gaps 1;

QY 13 EPRRTGEGVGPVGEVEMVKGQPFVDPVRYTQYIGEGAYGMVSSAYDHRKTRVAIKK 72
Db 1 EPRGTAGVVPVVEVGVVKGQPFVDPVRYTQYIGEGAYGMVSSAYDHRKTRVAIKK 60

QY 73 ISPFHEQTYCORTLREIQILLRFRHENVIGIRDLIRAPTEARMDVYIVQDLMETDLYKL 132
Db 61 ISPFHEQTYCORTLREIQILLRFRHENVIGIRDLIRAPTEARMDVYIVQDLMETDLYKL 120

QY 133 LKSQLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLXICDFGLARIAD 192
Db 121 LKSQLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLXICDFGLARIAD 180

QY 193 PEHDHTGFTYVATRWYRAPEIMLSNKGYSIDISVGCILAEMLSNRPFGKHLYLD 252
Db 181 PEHDHTGFTYVATRWYRAPEIMLSNKGYSIDISVGCILAEMLSNRPFGKHLYLD 240

QY 253 QLNHIL-----ALDLDRLMT 268
Db 241 QLNHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLDRLMT 300

QY 269 FNPKNRITVEEALAHPLYEQYDPTDEPVAERPPFTFAMELDDLPKRLKELIFOETARFQ 328
Db 301 FNPKNRITVEEALAHPLYEQYDPTDEPVAERPPFTFAMELDDLPKRLKELIFOETARFQ 360

QY 329 PGVLEAP 335
Db 361 PGAPAP 367

RESULT 10
US-08-461-985-11
; Sequence 11, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leellie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:

```

```
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-461-985-11

Query Match          91.2%; Score 1607; DB 1; Length 367;
Best Local Similarity 85.3%; Pred. No. 1.3e-160;
Matches 313; Conservative 1; Mismatches 9; Indels 44; Gaps 1;

QY 13 EPRRTGEGVGPGEVEMVKGQPFVGPRTQIQYIGEGAGVMSAYDHRKTRVAIKK 72
DB 1 EPRGTAGVVVPGVEVEMVKGQPFVGPRTQIQYIGEGAGVMSAYDHRKTRVAIKK 60
QY 73 ISPFHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKL 132
DB 61 ISPFHQTYCORTLREIQILLRFRHENVIGIRDILRAPTLEAMRDVYIVQDLMETDLYKL 120
QY 133 LKSQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIAD 192
DB 121 LKSQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIAD 180
QY 193 PEHDTGFTLEVYATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLD 252
DB 181 PEHDTGFTLEVYATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLD 240
QY 253 QLNHIL-----ALDLDRLMT 268
DB 241 QLNHILGLSPSQEDLNCIINNKARNYLOSPLSKTKVAVAKLFPKSDSKALDLDRLMT 300
QY 269 FNPKNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQ 328
DB 301 FNPKNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQ 360
QY 329 PGVLEAP 335
DB 361 PGAPAP 367

RESULT 11
US-08-459-953A-8
; Sequence 8, Application US/08459953A
; Patent No. 6030822
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,953A
; FILING DATE: June 2, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,494
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-953A-8

Query Match          91.1%; Score 1606; DB 2; Length 355;
Best Local Similarity 87.6%; Pred. No. 1.6e-160;
Matches 311; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 25 PGEVEMVKGQPDVGPRTQIQYIGEGAGVMSAYDHRKTRVAIKKISPFHQTYCOR 84
DB 1 PGEVEMVKGQPDVGPRTQIQYIGEGAGVMSAYDHRKTRVAIKKISPFHQTYCOR 60
QY 85 TLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLKSQLSNDHIC 144
DB 61 TLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLKSQLSNDHIC 120
QY 145 YPLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFTLEY 204
DB 121 YPLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFTLEY 180
QY 205 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQNLHIL----- 258
DB 181 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQNLHILGLSP 240
QY 259 -----ALDLDRLMTFNPKNRITVEEA 280
DB 241 SQEDLNCIINNKARNYLOSPLSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRITVEEA 300
QY 281 LAHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQGVLEAP 335
DB 301 LAHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQGVLEAP 355

RESULT 12
US-09-393-212-8
; Sequence 8, Application US/09393212
; Patent No. 6579972
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/393,212
FILING DATE: 09-Sep-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/459,953A
FILING DATE: June 2, 1995
APPLICATION NUMBER: 08/029,494
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-393-212-8

Query Match 91.1%; Score 1606; DB 2; Length 355;
Best Local Similarity 87.6%; Pred. No. 1.6e-160;
Matches 311; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 25 PGEVWVGQPPDVGPRYTQVIGEGAYGWSAYDHRVKTVAIKKISPPHEQTYCQR 84
DB 1 PGEVWVGQPPDVGPRYTQVIGEGAYGWSAYDHRVKTVAIKKISPPHEQTYCQR 60
QY 85 TLREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSLNDHIC 144
DB 61 TLREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSLNDHIC 120
QY 145 YFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFLTEY 204
DB 121 YFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFLTEY 180
QY 205 VATRWYRAPIMLNSKGYTYSIDWSVGCILAEMLSNRPIFPKGKHYLDQLNHIL----- 258
DB 181 VATRWYRAPIMLNSKGYTYSIDWSVGCILAEMLSNRPIFPKGKHYLDQLNHILGILGSP 240
QY 259 -----ALDLDRLMTFNPKNRITVEEA 280
DB 241 SQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSDSKALDLDRLMTFNPKNRITVEEA 300
QY 281 LAHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFPQGVLEAP 335
DB 301 LAHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFPQGVLEAP 355

RESULT 13
US-08-176-620A-2
Sequence 2, Application US/08176620A
Patent No. 5595904
GENERAL INFORMATION:
APPLICANT: Boulton, Teri G.
APPLICANT: Cobb, Melanie H.
APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayotatos, Nikos
TITLE OF INVENTION: A family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,620A
FILING DATE: 03-JAN-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Lealie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-620A-2

Query Match 90.5%; Score 1595; DB 1; Length 365;
Best Local Similarity 85.2%; Pred. No. 2.4e-159;
Matches 311; Conservative 1; Mismatches 9; Indels 44; Gaps 1;
QY 15 RTEGVGPGVCGEVEWVGQPPDVGPRYTQVIGEGAYGWSAYDHRVKTVAIKKIS 74
DB 1 RGTAGVVPVPGVEVVGQPPDVGPRYTQVIGEGAYGWSAYDHRVKTVAIKKIS 60
QY 75 PFEHQTYCQRTLEIREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLK 134
DB 61 PFEHQTYCQRTLEIREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLK 120
QY 135 SQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPE 194
DB 121 SQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPE 180
QY 195 HDHTGFLTEVATRWYRAPIMLNSKGYTYSIDWSVGCILAEMLSNRPIFPKGKHYLDQL 254
DB 181 HDHTGFLTEVATRWYRAPIMLNSKGYTYSIDWSVGCILAEMLSNRPIFPKGKHYLDQL 240
QY 255 NHIL-----ALDLDRLMTFNPKNRITVEEA 270
DB 241 NHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSDSKALDLDRLMTFNP 300
QY 271 PNRKITVEEALAHPIYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFPQGV 330
DB 301 PNRKITVEEALAHPIYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFPQGV 360
QY 331 VLEAP 335
DB 361 APEAP 365

RESULT 14
US-08-463-862-2
Sequence 2, Application US/08463862
Patent No. 5776751
GENERAL INFORMATION:
APPLICANT: Boulton, Teri G. et al.
TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,862
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,544
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-049
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-862-2

Query Match 90.5%; Score 1595; DB 1; Length 365;
Best Local Similarity 85.2%; Pred. No. 2.4e-159;
Matches 311; Conservative 1; Mismatches 9; Indels 44; Gaps 1;

QY	15	RTGEGVGPVGEVEMVKQPDVGPRTQIQYIGEGAYGMVSSAYDHRKTRVAIKKIS	74
DB	1	RGTAGVVPVPPGEVGVKQPDVGPRTQIQYIGEGAYGMVSSAYDHRKTRVAIKKIS	60
QY	75	PPEHTQYQORTUREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK	134
DB	61	PPEHTQYQORTUREIQILLGFRHENVIGIRDILRAFTLEAMRDVYIVQDLMETDLYKLLK	120
QY	135	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	194
DB	121	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	180
QY	195	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLSNRPFPKHYLDOL	254
DB	181	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLSNRPFPKHYLDOL	240
QY	255	NHIL-----ALDLDRLMTFN	270
DB	241	NHILGLSPSQEDLNCINMKARNYLSLPSTKVAVAKLFPKSDSKALDLDRLMTFN	300
QY	271	PNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPKRLKELIFQETARFQPG	330
DB	301	PNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPKRLKELIFQETARFQPG	360
QY	331	VLEAP 335	
DB	361	APEAP 365	

RESULT 15

US-08-461-985-2
; Sequence 2, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.

APPLICANT: Cobb, Melanie H.
APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayotatos, Nikos
TITLE OF INVENTION: A Family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,985
FILING DATE: 05-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,620
FILING DATE: 03-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-985-2

Query Match 90.5%; Score 1595; DB 1; Length 365;

Best Local Similarity 85.2%; Pred. No. 2.4e-159;

Matches 311; Conservative 1; Mismatches 9; Indels 44; Gaps 1;

QY	15	RTGEGVGPVGEVEMVKQPDVGPRTQIQYIGEGAYGMVSSAYDHRKTRVAIKKIS	74
DB	1	RGTAGVVPVPPGEVGVKQPDVGPRTQIQYIGEGAYGMVSSAYDHRKTRVAIKKIS	60
QY	75	PPEHTQYQORTUREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK	134
DB	61	PPEHTQYQORTUREIQILLGFRHENVIGIRDILRAFTLEAMRDVYIVQDLMETDLYKLLK	120
QY	135	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	194
DB	121	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	180
QY	195	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLSNRPFPKHYLDOL	254
DB	181	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLSNRPFPKHYLDOL	240
QY	255	NHIL-----ALDLDRLMTFN	270
DB	241	NHILGLSPSQEDLNCINMKARNYLSLPSTKVAVAKLFPKSDSKALDLDRLMTFN	300
QY	271	PNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPKRLKELIFQETARFQPG	330
DB	301	PNKRITVEEALAHPLYEQYDPTDFVABEPTTFAMELDDLPKRLKELIFQETARFQPG	360
QY	331	VLEAP 335	
DB	361	APEAP 365	

Search completed: February 6, 2006, 15:36:48
Job time : 34.7896 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:35:46 ; Search time 116.816 Seconds
(without alignments)
1198.238 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAQQGGGGPRTEGV.....LKELIFQETARFQPGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA_Main:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:*

2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:*

3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap:*

4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap:*

5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap:*

6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	335	5	US-10-623-108-2
2	1731	98.2	379	4	US-10-233-448-7
3	1731	98.2	379	4	US-10-174-794-14
4	1731	98.2	379	4	US-10-394-322A-19
5	1731	98.2	379	4	US-10-701-490-8
6	1731	98.2	379	5	US-10-735-118-3
7	1731	98.2	379	5	US-10-623-108-6
8	1731	98.2	379	5	US-10-945-684-2
9	1725	97.8	379	4	US-10-060-065-16
10	1725	97.8	379	4	US-10-059-585-37
11	1725	97.8	379	4	US-10-170-663-25
12	1725	97.8	379	5	US-10-893-072-25
13	1725	97.8	631	4	US-10-072-036-39
14	1659	94.1	624	4	US-10-072-036-57
15	1606	91.1	355	4	US-10-461-402-8
16	1599	90.7	359	5	US-10-623-108-4
17	1599	90.7	359	5	US-10-623-108-8
18	1599	90.4	359	4	US-10-114-270-110
19	1455.5	82.6	392	5	US-10-732-923-1420
20	1454.5	82.5	380	4	US-10-461-402-9
21	1444.5	81.9	360	4	US-10-171-311-127
22	1444.5	81.9	360	4	US-10-233-448-8
23	1444.5	81.9	360	4	US-10-301-822-115
24	1444.5	81.9	360	4	US-10-394-322A-20
25	1444.5	81.9	360	5	US-10-941-635-56
26	1438.5	81.6	360	4	US-10-664-421-56
27	1434.5	81.4	358	5	US-10-945-684-1

ALIGNMENTS

RESULT 1

US-10-623-108-2

; Sequence 2, Application US/10623108

; Publication No. US20050013817A1

; GENERAL INFORMATION:

; APPLICANT: DAI, KEN-SHMO

; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS

; FILE REFERENCE: U 014726-8

; CURRENT APPLICATION NUMBER: US/10/623,108

; CURRENT FILING DATE: 2003-07-18

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 335

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-623-108-2

Query Match 100.0%; Score 1763; DB 5; Length 335;

Best Local Similarity 100.0%; Pred. No. 1e-136; Indels 0; Gaps 0;

Matches 335; Conservative 0; Mismatches 0;

Qy 1 MAAAAAQQGGGGPRTEGVGPGEVEMVKQPFDPVGPRTYQLOVIGEGAYGMSSAY 60

Db 1 MAAAAAQQGGGGPRTEGVGPGEVEMVKQPFDPVGPRTYQLOVIGEGAYGMSSAY 60

Qy 61 DHVRKTRVAIKKISPFHQYTCQRTLRREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120

Db 61 DHVRKTRVAIKKISPFHQYTCQRTLRREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120

Qy 121 VQDLMETDLYKLKSKQSLNDHICYFLYQILRGKYIHSANVLRDLKPSNLLINTTCDL 180

Db 121 VQDLMETDLYKLKSKQSLNDHICYFLYQILRGKYIHSANVLRDLKPSNLLINTTCDL 180

Qy 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

Db 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

Qy 241 NRPIFFGKHVLDQNLHILADLLDRMLTFPNKRITVEEALAHPLYEQYDPTDPEVAEE 300

Db 241 NRPIFFGKHVLDQNLHILADLLDRMLTFPNKRITVEEALAHPLYEQYDPTDPEVAEE 300

Qy 301 PFTFAMELDDLPKRLKELIFQETARFQPGVLEAP 335

Db 301 PFTFAMELDDLPKRLKELIFQETARFQPGVLEAP 335

RESULT 2

US-10-233-448-7

; Sequence 7, Application US/10233448

```
; Publication No. US20030109419A1
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; TITLE OF INVENTION: METHOD FOR CLASSIFICATION OF ANTI-PSYCHOTIC DRUGS
; FILE REFERENCE: 11181-009
; CURRENT APPLICATION NUMBER: US/10/233,448
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 60/316,338
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 379
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-233-448-7

Query Match      98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAQGGGGGPRRTGPGVPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAQGGGGGPRRTGPGVPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
QY 61 DHVTRTRVAIKKISPEHQTYCORTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVTRTRVAIKKISPEHQTYCORTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
QY 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
QY 241 NRPIFGKHLYDLQNLHIL----- 258
   |||||
Db 241 NRPIFGKHLYDLQNLHILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLFPKSD 300
   |||||
QY 259 --ALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 316
   |||||
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 360
   |||||
QY 317 KELIFOETARFQPGVLEAP 335
   |||||
Db 361 KELIFOETARFQPGVLEAP 379
   |||||

RESULT 3
US-10-174-794-14
; Sequence 14, Application US/10174794
; Publication No. US20030166220A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PR1
; ORGANISM: Rattus norvegicus
; US-10-174-794-14
```

```
Query Match      98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAQGGGGGPRRTGPGVPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAQGGGGGPRRTGPGVPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
QY 61 DHVTRTRVAIKKISPEHQTYCORTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVTRTRVAIKKISPEHQTYCORTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
QY 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
QY 241 NRPIFGKHLYDLQNLHIL----- 258
   |||||
Db 241 NRPIFGKHLYDLQNLHILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLFPKSD 300
   |||||
QY 259 --ALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 316
   |||||
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKERL 360
   |||||
QY 317 KELIFOETARFQPGVLEAP 335
   |||||
Db 361 KELIFOETARFQPGVLEAP 379
   |||||

RESULT 4
US-10-394-322A-19
; Sequence 19, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 379
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-394-322A-19

Query Match      98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAQGGGGGPRRTGPGVPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAQGGGGGPRRTGPGVPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
QY 61 DHVTRTRVAIKKISPEHQTYCORTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVTRTRVAIKKISPEHQTYCORTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
QY 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
```

Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Qy 241 NRPIFGKHVLDQNLHIL----- 258
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVWAKLPKSD 300
Qy 259 --ALDILLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVASEPPTFAMELDDLPKRL 316
Db 301 SKALDILLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVASEPPTFAMELDDLPKRL 360
Qy 317 KELIFQETARFQGVLEAP 335
Db 361 KELIFQETARFQGVLEAP 379

RESULT 5

US-10-701-490-8
; Sequence 8, Application US/10701490
; Publication No. US20040106141A1
; GENERAL INFORMATION:
; APPLICANT: PAUL S. MISCHER
; APPLICANT: CHARLES L. SAWYERS
; APPLICANT: BRADLEY L. SMITH
; APPLICANT: KATHERINE CROSEY
; TITLE OF INVENTION: METHODS AND MATERIALS FOR EXAMINING
; FILE REFERENCE: PATHWAYS ASSOCIATED WITH GLOBLASTOMA PROGRESSION
; CURRENT APPLICATION NUMBER: US/10/701,490
; CURRENT FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/423,777
; PRIOR FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-701-490-8

Query Match 98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

Qy 1 MAAAAAGGGGGPRTEGVGPGEVEMVKGPDPVGPRTYQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVGPGEVEMVKGPDPVGPRTYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLKYLKTHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLKYLKTHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Qy 241 NRPIFGKHVLDQNLHIL----- 258
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVWAKLPKSD 300
Qy 259 --ALDILLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVASEPPTFAMELDDLPKRL 316
Db 301 SKALDILLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVASEPPTFAMELDDLPKRL 360
Qy 317 KELIFQETARFQGVLEAP 335
Db 361 KELIFQETARFQGVLEAP 379

RESULT 6

US-10-735-118-3
; Sequence 3, Application US/10735118
; Publication No. US20040248151A1
; GENERAL INFORMATION:
; APPLICANT: Bacus, Sarah S.
; APPLICANT: Smith, Bradley L.
; TITLE OF INVENTION: METHOD FOR PREDICTING THE RESPONSE TO HER2-DIRECTED THERAPY
; FILE REFERENCE: 6270-701.201
; CURRENT APPLICATION NUMBER: US/10/735,118
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 60/370,473
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/432,943
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 10/408,520
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-735-118-3

Query Match 98.2%; Score 1731; DB 5; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
Qy 1 MAAAAAGGGGGPRTEGVGPGEVEMVKGPDPVGPRTYQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVGPGEVEMVKGPDPVGPRTYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLKYLKTHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLKYLKTHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Qy 241 NRPIFGKHVLDQNLHIL----- 258
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVWAKLPKSD 300
Qy 259 --ALDILLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVASEPPTFAMELDDLPKRL 316
Db 301 SKALDILLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVASEPPTFAMELDDLPKRL 360
Qy 317 KELIFQETARFQGVLEAP 335
Db 361 KELIFQETARFQGVLEAP 379

RESULT 7

US-10-623-108-6
; Sequence 6, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHWO
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-623-108-6		US-10-623-108-2	
Query Match	98.2%; Score 1731; DB 5; Length 379;	Query Match	97.8%; Score 1725; DB 4; Length 379;
Best Local Similarity	88.4%; Pred. No. 5.2e-134;	Best Local Similarity	88.1%; Pred. No. 1.6e-133;
Matches 335; Conservative	0; Mismatches 0; Indels 44; Gaps 1;	Matches 334; Conservative	0; Mismatches 1; Indels 44; Gaps 1;
QY	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60	QY	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60
Db	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60	Db	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60
QY	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	QY	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	Db	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	QY	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	Db	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240	QY	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Db	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240	Db	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
QY	241 NRPIPGKHYLDQLNHIL----- 258	QY	241 NRPIPGKHYLDQLNHIL----- 258
Db	241 NRPIPGKHYLDQLNHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAVAKLPKSD 300	Db	241 NRPIPGKHYLDQLNHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAVAKLPKSD 300
QY	259 --ALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEERTFAMELDDLPKERL 316	QY	259 --ALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEERTFAMELDDLPKERL 316
Db	301 SKALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEERTFAMELDDLPKERL 360	Db	301 SKALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEERTFAMELDDLPKERL 360
QY	317 KELIFQETARFQGVLEAP 335	QY	317 KELIFQETARFQGVLEAP 335
Db	361 KELIFQETARFQGVLEAP 379	Db	361 KELIFQETARFQGVLEAP 379
RESULT 8		RESULT 9	
US-10-945-684-2		US-10-060-065-16	
; Sequence 2, Application US/10945684		; Sequence 16, Application US/10060065	
; Publication No. US2005009567A1		; Publication No. US20030017480A1	
; GENERAL INFORMATION:		; GENERAL INFORMATION:	
; APPLICANT: Emory University		; APPLICANT: Toshio Ota	
; TITLE OF INVENTION: Methods and Kits for Detecting Proteins		; APPLICANT: Takao Isogai	
; FILE REFERENCE: 50508-2280		; APPLICANT: Tetsuo Nishikawa	
; CURRENT APPLICATION NUMBER: US/10/945,684		; APPLICANT: Koji Hayashi	
; CURRENT FILING DATE: 2004-09-21		; APPLICANT: Kaoru Otsuka	
; NUMBER OF SEQ ID NOS: 8		; APPLICANT: Jun-Ichi Yamamoto	
; SOFTWARE: PatentIn version 3.2		; APPLICANT: Shizuko Ishii	
; SEQ ID NO 2		; APPLICANT: Tomoyasu Sugiyama	
; LENGTH: 379		; APPLICANT: Ai Wakamatsu	
; TYPE: PRT		; APPLICANT: Keiichi Negai	
; ORGANISM: Homo Sapiens		; APPLICANT: Tetsuji Otsuki	
US-10-945-684-2		; APPLICANT: Shin-Ichi Funahashi	
Query Match	98.2%; Score 1731; DB 5; Length 379;	; APPLICANT: Chiaki Senoo	
Best Local Similarity	88.4%; Pred. No. 5.2e-134;	; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE	
Matches 335; Conservative	0; Mismatches 0; Indels 44; Gaps 1;	; FILE REFERENCE: 06501-099002	
QY	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60	; CURRENT APPLICATION NUMBER: US/10/060,065	
Db	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60	; CURRENT FILING DATE: 2002-01-29	
QY	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	; PRIOR APPLICATION NUMBER: PCT/JP00/05061	
Db	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	; PRIOR FILING DATE: 2000-07-28	
QY	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	; PRIOR APPLICATION NUMBER: US 60/159,590	
Db	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	; PRIOR FILING DATE: 1999-10-18	
QY	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240	; PRIOR APPLICATION NUMBER: US 60/183,322	
Db	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240	; PRIOR FILING DATE: 2000-02-17	
QY	241 NRPIPGKHYLDQLNHIL----- 258	; PRIOR APPLICATION NUMBER: JP 11-248036	
Db	241 NRPIPGKHYLDQLNHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAVAKLPKSD 300	; PRIOR FILING DATE: 1999-07-29	
QY	259 --ALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEERTFAMELDDLPKERL 316	; PRIOR APPLICATION NUMBER: JP 2000-118776	
Db	301 SKALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEERTFAMELDDLPKERL 360	; PRIOR FILING DATE: 2000-01-11	
QY	317 KELIFQETARFQGVLEAP 335	; PRIOR APPLICATION NUMBER: JP 2000-183767	
Db	361 KELIFQETARFQGVLEAP 379	; PRIOR FILING DATE: 2000-05-02	
US-10-945-684-2		; PRIOR APPLICATION NUMBER: JP 2000-241899	
Query Match	98.2%; Score 1731; DB 5; Length 379;	; NUMBER OF SEQ ID NOS: 43	
Best Local Similarity	88.4%; Pred. No. 5.2e-134;	; SOFTWARE: PatentIn Ver. 2.0	
Matches 335; Conservative	0; Mismatches 0; Indels 44; Gaps 1;	; SEQ ID NO 16	
QY	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60	; LENGTH: 379	
Db	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60	; TYPE: PRT	
QY	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	; ORGANISM: Homo sapiens	
Db	61 DHVRKTRVAIKKISPEPHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120	US-10-060-065-16	
QY	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	Query Match	97.8%; Score 1725; DB 4; Length 379;
Db	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	Best Local Similarity	88.1%; Pred. No. 1.6e-133;
QY	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240	Matches 334; Conservative	0; Mismatches 1; Indels 44; Gaps 1;
Db	181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240	QY	1 MAAAAQGGGGEPRTGEGVGPVGEVEMVKGQPPDVGPRTYQIYIGEGAYGMVSSAY 60

QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
 DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
 QY 241 NRPIFGKHYLDQNLHIL-----258
 DB 241 NRPIFGKHYLDQNLHIL-----258
 QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 316
 DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 360
 QY 317 KELIFQETARFQGVLEAP 335
 DB 361 KELIFQETARFQGVLEAP 379

RESULT 10
 US-10-059-585-37
 ; Sequence 37, Application US/10059585
 ; Publication No. US20030082776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ota, Toshio
 ; APPLICANT: Isogai, Takao
 ; APPLICANT: Nishikawa, Tetsuo
 ; APPLICANT: Hayashi, Koji
 ; APPLICANT: Otsuka, Kaoru
 ; APPLICANT: Yamamoto, Jun-ichi
 ; APPLICANT: Ishii, Shizuko
 ; APPLICANT: Sugiyama, Tomoyasu
 ; APPLICANT: Wakamatsu, Ai
 ; APPLICANT: Nagai, Keiichi
 ; APPLICANT: Otsuki, Tetsuji
 ; APPLICANT: Funahashi, Shin-Ichi
 ; APPLICANT: Nezu, Jun-Ichi
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
 ; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
 ; FILE REFERENCE: 06501-098001
 ; CURRENT APPLICATION NUMBER: US/10/059,585
 ; CURRENT FILING DATE: 2002-01-29
 ; PRIOR APPLICATION NUMBER: PCT/JP00/05060
 ; PRIOR FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: US 60/183,322
 ; PRIOR FILING DATE: 2000-02-17
 ; PRIOR APPLICATION NUMBER: US 60/159,590
 ; PRIOR FILING DATE: 1999-10-18
 ; PRIOR APPLICATION NUMBER: JP 2000-118776
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: JP 2000-183767
 ; PRIOR FILING DATE: 2000-05-02
 ; PRIOR APPLICATION NUMBER: JP 11-248036
 ; PRIOR FILING DATE: 1999-07-29
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 37
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-059-585-37

Query Match 97.8%; Score 1725; DB 4; Length 379;
 Best Local Similarity 88.1%; Pred. No. 1.6e-133;
 Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;
 QY 1 MAAAACGGGCGEPRRTGEGVGPVGEVEMVKGQPFDPGPRYTQIQVIGEGAYGMVSSAY 60
 DB 1 MAAAACGGGCGEPRRTGEGVGPVGEVEMVKGQPFDPGPRYTQIQVIGEGAYGMVSSAY 60
 QY 61 DHVAKTRVAIKKISPFHEQTQCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
 DB 61 DHVAKTRVAIKKISPFHEQTQCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDYKLKSOQLSNDHICYFLYQILRLGKLYIHSANVLHRLDKPSNLLINTTCDL 180
 DB 121 VQDLMETDYKLKSOQLSNDHICYFLYQILRLGKLYIHSANVLHRLDKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
 DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
 QY 241 NRPIFGKHYLDQNLHIL-----258
 DB 241 NRPIFGKHYLDQNLHIL-----258
 QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 316
 DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 360
 QY 317 KELIFQETARFQGVLEAP 335
 DB 361 KELIFQETARFQGVLEAP 379

RESULT 11
 US-10-170-663-25
 ; Sequence 25, Application US/10170663
 ; Publication No. US20030165899A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Su, Michael Shin-San
 ; Fox, Ted
 ; Wilson, Keith Phillip
 ; Germann, Ursula A.
 ; TITLE OF INVENTION: Methods For Designing Inhibitors of
 ; Serine/Threonine Kinases and Tyrosine Kinase
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Neave
 ; STREET: 1251 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: US
 ; ZIP: 10020
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/170,663
 ; FILING DATE: 12-Jun-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/025,580
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haley, James F.
 ; REGISTRATION NUMBER: 27,794
 ; REFERENCE/DOCKET NUMBER: VPI 97-104
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 596-9000
 ; TELEFAX: (212) 596-9090
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 379 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
 US-10-170-663-25

Query Match 97.8%; Score 1725; DB 4; Length 379;
 Best Local Similarity 88.1%; Pred. No. 1.6e-133;
 Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

Db 253 MAAAAAQQGGGPRTEGVGPGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAY 121
 QY 61 DHVKTTRVAIKKISPEHQTYCORTLEIQILRRFHNENIGIRDIRASTLEAMRDVVI 120
 Db 313 DHVKTTRVAIKKISPEHQTYCORTLEIQILRRFHNENIGIRDIRASTLEAMRDVVI 372
 QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 Db 373 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 432
 QY 181 KICDFGLARIADPEHDHGTGLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
 Db 433 KICDFGLARIADPEHDHGTGLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 492
 QY 241 NRPIFGKHYLDQNLHIL----- 258
 Db 493 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSLPSKTKVAKLFPKSD 552
 QY 259 --ALDLDRLMTNPNKRIITVEBALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRL 316
 Db 553 SKALDLDRLMTNPNKRIITVEBALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRL 612
 QY 317 KELIFQETARFQGVLEAP 335
 Db 613 KELIFQETARFQGVLEAP 631

RESULT 14
 US-10-072-036-57
 ; Sequence 57, Application US/10072036
 ; Publication No. US20030082564A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ole THASTRUP
 ; APPLICANT: Sara BJRON
 ; APPLICANT: Soren TULLIN
 ; APPLICANT: Kasper ALMHOLT
 ; APPLICANT: Kurt SCUDDER
 ; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
 ; FILE REFERENCE: 3759-0120P
 ; CURRENT APPLICATION NUMBER: US/10/072.036
 ; PRIOR FILING DATE: 2002-09-13
 ; PRIOR FILING DATE: 1999-10-07
 ; NUMBER OF SEQ ID NOS: 143
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 57
 ; LENGTH: 624
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: MERK1-F64L-S65T-GFP fusion
 US-10-072-036-57

Query Match 94.1%; Score 1659; DB 4; Length 624;
 Best Local Similarity 85.4%; Pred. No. 8e-128;
 Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;
 QY 2 AAAAAQQGGGPRTEGVGPGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAYD 61
 Db 3 AAAAAQQGGGPRTEGVGPGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAYD 62
 QY 62 HVRKTRVAIKKISPEHQTYCORTLEIQILRRFHNENIGIRDIRASTLEAMRDVVI 121
 Db 63 HVRKTRVAIKKISPEHQTYCORTLEIQILRRFHNENIGIRDIRASTLEAMRDVVI 122
 QY 122 QDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK 181
 Db 123 QDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK 182
 QY 182 KICDFGLARIADPEHDHGTGLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN 241
 Db 183 KICDFGLARIADPEHDHGTGLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN 242

QY 242 RPIFGKHYLDQNLHIL----- 258
 Db 243 RPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSLPSKTKVAKLFPKSDS 302
 QY 259 -ALDLDRLMTNPNKRIITVEBALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRLK 317
 Db 303 KALDLDRLMTNPNKRIITVEBALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRLK 362
 QY 318 ELIFQETARFQGVLEAP 335
 Db 363 ELIFQETARFQGAPEGP 380

RESULT 15
 US-10-461-402-8
 ; Sequence 8, Application US/10461402
 ; Publication No. US20030229209A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lechner, Cornelia
 ; Moller, Niels P.H.
 ; Ullrich, Axel
 ; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
 ; KINASE, SEQUENCES, AND
 ; METHODS OF PRODUCTION
 ; AND USE
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071-2066
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0
 ; SOFTWARE: Microsoft Word
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/461,402
 ; FILING DATE: 16-Jun-2003
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/459,953A
 ; FILING DATE: June 2, 1995
 ; APPLICATION NUMBER: 08/029,494
 ; FILING DATE: March 19, 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warburg, Richard J.
 ; REGISTRATION NUMBER: 32,327
 ; REFERENCE/DOCKET NUMBER: 209/267
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 355 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-10-461-402-8

Query Match 91.1%; Score 1606; DB 4; Length 355;
 Best Local Similarity 87.6%; Pred. No. 9.4e-124;
 Matches 311; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 25 PGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAYDHRKTRVAIKKISPEHQTYCOR 84

Db	1	PGEVEMVKQPFDVGPRTQLOVIGBAYGMVSSAYDHVRKTRVAIKKISPPHEQTYCQR	60
Qy	85	TLREIQIILLRFRHENVIGIRDIIRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC	144
Db	61	TLREIQIILLRFRHENVIGIRDIIRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC	120
Qy	145	YFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLTEY	204
Db	121	YFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLTEY	180
Qy	205	VATRWYRAPEIMLNSKGYTKSIDINSVGCILAEMLSNRPIFFGKHYLDQLNHIL-----	258
Db	181	VATRWYRAPEIMLNSKGYTKSIDINSVGCILAEMLSNRPIFFGKHYLDQLNHILGILGSP	240
Qy	259	-----ALDLDRLMTFNPKNRITYVEA	280
Db	241	SOEDLNCIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRITYVEA	300
Qy	281	LAHPYLEQYVDPTDEPVASEPPTFAMELDDLPKERLKELI FORTARFQGVLEAP	335
Db	301	LAHPYLEQYVDPTDEPVASEPPTFAMELDDLPKERLKELI FORTARFQGVLEAP	355

Search completed: February 6, 2006, 15:40:56
Job time : 117.816 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:37:06 ; Search time 10.1369 Seconds
(without alignments)
387.254 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAGGGGGPRTEGV.....LKELIFQETARFGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1725	97.8	379	7	US-11-109-156-16	Sequence 16, Appl
2	1444.5	81.9	360	6	US-10-878-556A-134	Sequence 134, App
3	1444.5	81.9	360	7	US-11-186-284-115	Sequence 115, App
4	747.5	42.4	360	6	US-10-886-329-1	Sequence 1, Appli
5	678.5	38.5	365	6	US-10-770-726-69	Sequence 69, Appl
6	620	35.2	422	6	US-10-857-780-21	Sequence 21, Appl
7	620	35.2	422	7	US-11-127-817-18	Sequence 18, Appl
8	620	35.2	426	7	US-11-127-817-20	Sequence 20, Appl
9	620	35.2	464	7	US-11-127-817-19	Sequence 19, Appl
10	464.5	26.3	294	7	US-11-024-959-263	Sequence 263, App
11	464.5	26.3	302	7	US-11-024-959-262	Sequence 262, App
12	461	26.1	304	7	US-11-024-959-264	Sequence 264, App
13	460.5	26.1	294	7	US-11-024-959-265	Sequence 265, App
14	454.5	25.8	460	7	US-11-024-959-384	Sequence 384, App
15	450.5	25.6	520	7	US-11-024-959-272	Sequence 272, App
16	449	25.5	277	7	US-11-127-817-21	Sequence 21, Appl
17	447	25.4	298	6	US-10-770-726-51	Sequence 51, Appl
18	442	25.1	795	6	US-10-770-726-49	Sequence 49, Appl
19	434	24.6	292	6	US-10-770-726-53	Sequence 53, Appl
20	432.5	24.5	297	6	US-10-770-726-48	Sequence 48, Appl
21	432.5	24.5	297	7	US-11-109-156-11	Sequence 11, Appl
22	432.5	24.5	346	6	US-10-770-726-55	Sequence 55, Appl
23	427.5	24.2	463	7	US-11-024-959-395	Sequence 395, App
24	418.5	23.7	509	7	US-11-024-959-393	Sequence 393, App
25	414.5	23.5	555	7	US-11-024-959-270	Sequence 270, App

RESULT 1
US-11-109-156-16
; Sequence 16, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; PRIOR FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-16

ALIGNMENTS

Sequence 72, Appl
Sequence 183, App
Sequence 267, App
Sequence 72, Appl
Sequence 65, Appl
Sequence 385, App
Sequence 5, Appl
Sequence 271, App
Sequence 261, App
Sequence 12, App
Sequence 269, App
Sequence 8, Appl
Sequence 273, App
Sequence 9, Appl
Sequence 10, Appl
Sequence 268, App
Sequence 386, App
Sequence 391, App
Sequence 52, App
Sequence 266, App

26 413.5 23.5 256 6 US-10-877-346-72
27 413.5 23.5 256 7 US-11-113-424-183
28 411 23.3 313 7 US-11-024-959-267
29 410 23.3 496 6 US-10-770-726-72
30 409 23.2 433 6 US-10-770-726-65
31 408 23.1 476 7 US-11-024-959-385
32 406 23.0 435 6 US-10-786-065-5
33 402 22.8 469 7 US-11-024-959-271
34 401 22.7 477 7 US-11-024-959-261
35 401 22.7 483 6 US-10-451-375-12
36 399 22.6 706 7 US-11-024-959-269
37 397 22.5 330 6 US-10-786-065-8
38 396 22.5 718 7 US-11-024-959-273
39 395 22.4 330 6 US-10-786-065-10
40 395 22.4 330 6 US-10-786-065-9
41 394.5 22.4 583 7 US-11-024-959-268
42 394 22.3 302 7 US-11-024-959-386
43 389 22.1 845 7 US-11-024-959-391
44 387.5 22.0 303 6 US-10-770-726-52
45 381.5 21.6 305 7 US-11-024-959-266

```
Query Match      97.8%; Score 1725; DB 7; Length 379;
Best Local Similarity 88.1%; Pred. No. 5.3e-147;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAQGGGGPRRTGEGVGPGEVEMVKQPPDVGPRYTQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAQGGGGPRRTGEGVGPGEVEMVKQPPDVGPRYTQLOYIGEGAYGMVSSAY 60
   |||||

QY 61 DHVRKTRVAIKKISPEHQTYCORTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVRKTRVAIKKISPEHQTYCORTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||

QY 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||

QY 241 NRPIPGKHLYDLQNLHIL----- 258
   |||||
Db 241 NRPIPGKHLYDLQNLHIL----- 258
   |||||

QY 259 --ALDILDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLKPERL 316
   |||||
Db 259 --ALDILDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLKPERL 316
   |||||

QY 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLKPERL 360
   |||||
Db 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLKPERL 360
   |||||

QY 317 KELIFOETARFQGVLEAP 335
   |||||
Db 361 KELIFOETARFQGVLEAP 379
   |||||

RESULT 2
US-10-878-556A-134
; Sequence 134, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/mx01_human
; DATABASE ENTRY DATE: 1992-12-01
US-10-878-556A-134

Query Match      81.9%; Score 1444.5; DB 6; Length 360;
Best Local Similarity 74.3%; Pred. No. 6.1e-122;
Matches 278; Conservative 18; Mismatches 17; Indels 61; Gaps 2;

QY 1 MAAAAQGGGGPRRTGEGVGPGEVEMVKQPPDVGPRYTQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAAGAG-----PEMVRQVDFVGPRTNLSYIGEGAYGMVCSAY 43
   |||||

QY 61 DHVRKTRVAIKKISPEHQTYCORTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 44 DNVNKRVAIKKISPEHQTYCORTREIKILLRFRHENIIGINDIIRAPTIEQMKDVYI 103
   |||||

QY 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
Db 104 VQDLMETDLYKLLKTQHLNSNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 163
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 164 KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 223
   |||||

QY 241 NRPIPGKHLYDLQNLHIL----- 258
   |||||
Db 224 NRPIPGKHLYDLQNLHIL----- 258
   |||||
```

```
QY 241 NRPIPGKHLYDLQNLHIL----- 258
   |||||
Db 224 NRPIPGKHLYDLQNLHIL----- 258
   |||||

QY 259 --ALDILDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLKPERL 316
   |||||
Db 284 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEPTTFAMELDDLKPERL 343
   |||||

QY 317 KELIFOETARFQGP 330
   |||||
Db 344 KELIFEETARFQGP 357
   |||||

RESULT 3
US-11-186-284-115
; Sequence 115, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-115

Query Match      81.9%; Score 1444.5; DB 7; Length 360;
Best Local Similarity 74.3%; Pred. No. 6.1e-122;
Matches 278; Conservative 18; Mismatches 17; Indels 61; Gaps 2;

QY 1 MAAAAQGGGGPRRTGEGVGPGEVEMVKQPPDVGPRYTQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAAGAG-----PEMVRQVDFVGPRTNLSYIGEGAYGMVCSAY 43
   |||||

QY 61 DHVRKTRVAIKKISPEHQTYCORTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 44 DNVNKRVAIKKISPEHQTYCORTREIKILLRFRHENIIGINDIIRAPTIEQMKDVYI 103
   |||||

QY 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
   |||||
Db 104 VQDLMETDLYKLLKTQHLNSNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 163
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 164 KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 223
   |||||

QY 241 NRPIPGKHLYDLQNLHIL----- 258
   |||||
Db 224 NRPIPGKHLYDLQNLHIL----- 258
   |||||
```

```
Qy 259 --ALDLDRLMTNPNKRTVEBALAHPLYEQYDTPDEPVAREPPTFAMELDDLPKRL 316
Db 284 SKALDLDKMLTNPKHRIEVEQALAHPLYEQYDTPDEPVAREPPTFAMELDDLPKRL 343
Qy 317 KELIFOETARFPG 330
Db 344 KELIFETARFPG 357

RESULT 4
US-10-886-329-1
; Sequence 1, Application US/10886329
; Publication No. US20050288286A1
; GENERAL INFORMATION:
; APPLICANT: Deciphra Pharmaceuticals, Inc.
; APPLICANT: Flynn, Daniel L
; APPLICANT: Petillo, Peter A
; TITLE OF INVENTION: Anti-Inflammatory Medicaments
; FILE REFERENCE: 34477-CIP
; CURRENT APPLICATION NUMBER: US/10/886,329
; CURRENT FILING DATE: 2004-07-06
; PRIOR FILING DATE: 2004-12-31
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US 10/746,460
; PRIOR FILING DATE: 2003-12-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; TYPE: PRT
; LENGTH: 360
; ORGANISM: Homo sapiens
US-10-886-329-1

Query Match 42.4%; Score 747.5; DB 6; Length 360;
Best Local Similarity 45.1%; Pred. No. 1.3e-59;
Matches 155; Conservative 55; Mismatches 81; Indels 53; Gaps 6;

Qy 36 FDVGPRTQYIGEGAYGMVSSAYDHRKTRVAIKKIS-PFEHQTYCORTLREIQILLR 94
Db 18 WEVPERYQNLSPVSGAYGVCVCAAFPTKTLRVAVKLSRPFQSIHAKRTYRELRLKH 77
Qy 95 FRHENVIGIRDLR-ASTLEAMRDVYVQDLMETDLYKLLKSQOLSDNHCYFLYQILRG 153
Db 78 MGHENVIGLDDVFTPARSLSEDFNDVYLVTLHMGADLNIVKQCKLTDHQQFLYQILRG 137
Qy 154 LKVIHSANVLHRLKPSNLINTTCDLKI CDFGLARIADPEHDHTGFLTEYVATRWYRAP 213
Db 138 LKVIHSADIIHRDKPSNLAVNEDCELKILDFGLARHTDDE-----MTGYVATRWYRAP 191
Qy 214 EIMLSNGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQL----- 254
Db 192 EIMLNNHYNQTVDIWSVGCIMAEMLTGTTLFPCTDHDIDQLKILRLVGTGPAELLKKIS 251
Qy 255 -----NHI-----LALDLDRLMTNPNKRTVEBALAHPLYEQY 289
Db 252 SESARNYIQSLTPKMPKNFANVFICANPLAVDLLEKMLVLDSDKRTITAAQALAHAYPAQY 311
Qy 290 YDPTDEPVAREPPTFAMELDDLPKRLKELIFOETARFPGVLE 333
Db 312 HDPDPEVA-DPYDQSPESRDLIDEWKSLTYDEVISFVPPPLD 354

RESULT 5
US-10-770-726-69
; Sequence 69, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)
```

```
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 69
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-69

Query Match 38.5%; Score 678.5; DB 6; Length 365;
Best Local Similarity 41.6%; Pred. No. 2e-53;
Matches 139; Conservative 55; Mismatches 87; Indels 53; Gaps 5;

Qy 42 YTQLOVIGEGAYGMVSSAYDHRKTRVAIKKIS-PFEHQTYCORTLREIQILLRPFHENV 100
Db 25 YVSPTHVSGAYGVCVCAIDKRSGEKVAIKLSRPFQSEIFAKRAYRELLLLKHMQHENV 84
Qy 101 IGIRDILR-ASTLEAMRDVYVQDLMETDLYKLLKSQOLSDNHCYFLYQILRLGLKYIHS 159
Db 85 IGLLDVFTTPASSLRNFYDLYVNPFPMTDLQKIM-GMEFSEBKIQYLVYQMLKGLKYIHS 143
Qy 160 ANVLHRLKPSNLINTTCDLKI CDFGLARIADPEHDHTGFLTEYVATRWYRAPPEIMLS 219
Db 144 AGVHEDLKPSNLAVNEDCELKILDFGLARHADA-----MTGYVATRWYRAPPEVILSW 197
Qy 220 KGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQL----- 258
Db 198 MHNQTVDIWSVGCIMAEMLTGTTLFPKGYLDQLTKILKVTGPGTGVPGTGVQKLNKAAKS 257
Qy 259 -----ALDLDRLMTNPNKRTVEBALAHPLYEQYDTPDE 295
Db 258 YIQSLPQTPRKDFPTQLFPRASQADLLEKMLVLDVDRKLTAAQALTHPFEFPRDPEE 317
Qy 296 PVAREPPTFAMELDDLPKRLKELIFOETARFOP 329
Db 318 TEAQPFDDSLSEKLTVDENKQHIYKEIVNFSF 351

RESULT 6
US-10-857-780-21
; Sequence 21, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-780-21

Query Match 35.2%; Score 620; DB 6; Length 422;
Best Local Similarity 38.2%; Pred. No. 4.2e-48;
Matches 133; Conservative 67; Mismatches 80; Indels 68; Gaps 9;
```

Qy	36	FVGPRTYQLYIGGAVGMSSADVHVRKTRVAIKKIS--PFHQYVCQRTLEIQILLR	94
Db	58	FTVLKRYQMLKPIGSGAQIICAAAYDALVDRNVAIKKLSRPFQNTHAKRAYVELVLMKC	117
Qy	95	FHEENVIGTRDILR--ASTLEAMRDVYIVODLWETDLYKLLKSQOOLSNDHICFYLIQILLG	153
Db	118	VNHKNIIISLNVFTPKTLEEFQDYLVMELMDANLCQVIQ--MELDHERMSYLLYQMLCG	176
Qy	154	LKVIHSANVLHRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGF--LTEYVATRWYRA	212
Db	177	IKHLSAGIIHRLDKPSNIWVKRSDCTLKILDFGLARTAG-----TSFMTPIYVTVTRIYRA	231
Qy	213	PEIMLNSKGYTKSIDISWSGVCITLAEMLNSRPIFPKGHYLDQLNHIL-----	258
Db	232	PEVIL--GMGYKENVDIWSGCIINGEWPHKILFFGRDYIDQWNKVIYEQLTGCPCEPKMKL	290
Qy	259	-----ALDLDLRMLTFNPNKRITV	277
Db	291	QPTVRVYENRPKYAGLTFPKLFPDLSFPADSBHNKLKASQARDLLSKMLVIDPAKRISV	350
Qy	278	ESALAHPLYEQYDPTDFEVAEEPTFAMELDDLPK--ERLKELITQE	323
Db	351	DDALOHPIYINWYDPA--EVEAPPPQIYDKOLDEREHTIEEWKELIYKE	397

RESULT 7

```

US-11-127-817-18
; Sequence 18, Application US/11127817
; Publication No. US20050287519A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koenraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions And Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27,800-D USA
; CURRENT APPLICATION NUMBER: US/11/127,817
; CURRENT FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 534
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 422
; TYPE: prt
; ORGANISM: Homo sapiens
US-11-127-817-18

```

Overall Match 35 28. Score 620. DB 7. Length 422.

Query Match	35.2%; Score 620; DB 7; Length 422;
Best Local Similarity	38.2%; Pred. No. 4.2e-48;
Matches 133; Conservative	67; Mismatches 80; Indels 68; Gaps 9;
Qy	36 FDVGPRTYQLYIGEGAYGWMVSSAYDHRKTRVAIKKIS-PFEHDTYCORTREIQILLR 94
Db	58 FTVLKRYQNLKPGSGAQGIVCAAYDAVDLBNVAIKLSRPFQNTAKAYRELVLKMC 117
Qy	95 FRHENVTGIRDIILR-ASTLEAMRDVYIVQDLMETDLYKLLKSOOLSNDHICYFPLYQILRG 153
Db	118 VNHKNIIISLLNVFTFPQKTLEEFQDLYVLMELMDANLCVQIQ-MELDHRSYLLYQWLGC 176
Qy	154 LKVIHSANVLHRDLKPSNLIINTCCDLKI CDPLGLARIADPEHDHTGF-LTEYVATRWYRA 212
Db	177 IKHLHSAGIIHRDLKPSNIVVKSDCTLKILDPLGLARTAG-----TSFMTPTVVVTRYRA 231
Qy	213 PETMLNSKGYTKSDIIVSGVCGILAEMI.SNRP1FGKHVLDQLNHIL----- 258
Db	232 PEVIL-GMGYKENVDIIVSGCIMGEMVRRHKILFFGRDYIDQWNKVIQLGTGTCPEFWKKL 290
Qy	259 -----ALDLLDRMLTTFNFKRITY 277

```

Db      291  OPTVRNVNVRPKVAGLTPKLFPPDSLFPADSEHNKLGASQARDLLSKMLVIDPAKREISV 350
Qy      278  EEALAHPIYLEQYVYDTPDEPVASEPPTFAMELDLPK--EELKELIQE 323
Db      351  DDALQHPYINNVYDPA-EVEAPPQIYDQLDEREHTIBEWKELLYKE 397

RESULT 8
US-11-127--817-20
: Sequence 20, Application US/11127817
: Publication No. US20050287519A1
: GENERAL INFORMATION:
: APPLICANT: Merchiers, Pascal G.
: APPLICANT: Hoffmann, Marcel
: APPLICANT: Spittaels, Koenraad F. F.
: APPLICANT: Laenen, Wendy
: TITLE OF INVENTION: Methods, Compositions And Compound Assays For Inhibiting
: TITLE OF INVENTION: Amyloid-Beta Protein Production
: FILE REFERENCE: P27,800-D USA
: CURRENT APPLICATION NUMBER: US/11/127,817

```

RESULT 9

; Sequence 19, Application US/11127817
 ; Publication No. US20050287519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Merchiers, Pascal G.
 ; APPLICANT: Hoffmann, Marcel
 ; APPLICANT: Spittaels, Koenraad F. F.
 ; APPLICANT: Laenen, Wendy
 ; TITLE OF INVENTION: Methods, Compositions And Compound Assays For Inhibiting
 ; TITLE OF INVENTION: Amyloid-Beta Protein Production

; FILE REFERENCE: P27,800-D USA
 ; CURRENT APPLICATION NUMBER: US/11/127,817
 ; CURRENT FILING DATE: 2005-05-12
 ; PRIOR APPLICATION NUMBER: 60/570,352
 ; PRIOR FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: 60/603,948
 ; PRIOR FILING DATE: 2004-08-24
 ; NUMBER OF SEQ ID NOS: 534
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 19
 ; LENGTH: 464
 ; TYPE: PR
 ; ORGANISM: Homo sapiens
 US-11-127-817-19
 Query Match 35.2%; Score 620; DB 7; Length 464;
 Best Local Similarity 38.2%; Pred. No. 4.8e-48;
 Matches 133; Conservative 67; Mismatches 80; Indels 68; Gaps 9;
 QY 36 FDVGPRTQYIGEGAGYMSAYDHRVTRVAIKKIS-PPEHQYTCQRTLEIQLLR 94
 DB 58 FTVLKYQNLKPTGSAQOIGCAAYDAVLDNRVNAIKKLSRPFQNTAKRAYRELVMKC 117
 QY 95 FRENHVGIRDLR-ASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHICVFLYQILRG 153
 DB 118 VNHKNISLNVPTPQKTEEFQDVLVWELMDANLCVQIQ-MELDHKMSYLIYQMLCG 176
 QY 154 LKTIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGF-LTEYVATRWYRA 212
 DB 177 IKHLBSAGIHRDLKPSNIVKSDCTLKILDFGLARTAG-TSFMTYVTVTRYRA 231
 QY 213 PEIMNSKGYTSIDISWVGCILAEMLSNRPIPPGKHVLDQNLHL----- 258
 DB 232 PEVIL-GMGYKENVDISWVGCINGEMVRHKILFPGRDYIDQWNVIEQLGTPCPPEFMKCL 290
 QY 259 -----ALDLDRMLTFNPNKRITV 277
 DB 291 QPTVRNVENRKYAGLTFFKLPFOSLFPADSEHNKLSKQALVIDPAKRISV 350
 QY 278 EEALAHPLYEQYDPTDEPVABEPFFAMELDOLPK--ERLKELI FOE 323
 DB 351 DDALQHPYINWTDPA-EVEAPPPQIYDQLODERHTIEEWKELIYKE 397
 RESULT 10
 US-11-024-959-263
 ; Sequence 263, Application US/11024959
 ; Publication No. US20060010516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FORSTER, RICHARD L.
 ; APPLICANT: CONNETT, MARIE B.
 ; APPLICANT: EMERSON, SARAH JANE
 ; APPLICANT: GRIGOR, MURRAY ROBERT
 ; APPLICANT: HIGGINS, COLLEEN M.
 ; APPLICANT: LUND, STEVEN TROY
 ; APPLICANT: MAGUSIN, ANDREAS
 ; APPLICANT: KODRZYCKI, BOB
 ; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
 ; FILE REFERENCE: 044463-0360
 ; CURRENT APPLICATION NUMBER: US/11/024,959
 ; CURRENT FILING DATE: 2004-12-30
 ; PRIOR APPLICATION NUMBER: 60/533,036
 ; PRIOR FILING DATE: 2003-12-30
 ; NUMBER OF SEQ ID NOS: 782
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 263
 ; LENGTH: 294
 ; TYPE: PR
 ; ORGANISM: Eucalyptus sp.
 US-11-024-959-263
 Query Match 26.3%; Score 464.5; DB 7; Length 294;
 Best Local Similarity 36.0%; Pred. No. 2.1e-34;
 Matches 105; Conservative 46; Mismatches 87; Indels 57; Gaps 6;
 QY 41 RYTOLOQYIGEGAGYMSAYDHRVTRVAIKKIS-PPEHQYTCQRTLEIQLLRFRHEN 99
 DB 3 KYOQLAKIGEGTYGIVYKAKDKSGELLALKKRLAEDEGIPSTAREISLLQLOHPN 62
 QY 100 VIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLKS---QOLSNDHICVFLYQILRG 156
 DB 63 IVPLYDVVHTE-----KKLTLVFEFLDQDLKYLACDGNGLPEYTVKSFYQLQGI 117
 QY 157 IHSANVLRDLKPSNLLINTTCDLKICDFGLAR-IADPEHDHTGFLTYVATRWYRAPEI 215
 DB 118 CHEHRVLRDLKPSNLLINMEGELKADFGARAFGI PVRYN-THVVVTLWYRAPDV 173
 QY 216 MLNSKGYTSIDISWVGCILAEMLSNRPIFPCKHYLDQNLHL----- 258
 DB 174 LMGRKYSTQVDISWVGCIFAEMVNGRPLFPGSSQDQLLRIFKTLGTPPSLKTWPGMAEL 233
 QY 259 -----ALDLDRMLTFNPNKRITVEEALAHPLYE 287
 DB 234 PDFKNFPKYVQSFKKICPKKLDKTGLDLSRLMLOYDPAKRISAQAMGHPYFK 288

Db 115 YMRNGYFHRDLKPENLLV--TSOLIKIADFGlareVLTSPY-----TDYVSTRWYRA 166
Qy 213 PEIMNSKGYTKSIDISVGCILAEMLSNRPIFPKGHYLDQLNHIL----- 258
Db 167 PEVLQSPYTTAIDMWAVGAILAELFTLHPLFGESELDDEIYKICGVLGTPDYETWPDG 226
Qy 259 -----ALDLLDRMLTFNPNKRITVEEALAHPY 285
Db 227 MQLAAFRNFIPQPLPVNLSVLIPHASPEAIDLITRLCSWDPOKRPPTAEQALHPF 282

RESULT 15

US-11-024-959-272
; Sequence 272, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 272
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-272

Query Match 25.6%; Score 450.5; DB 7; Length 520;
Best Local Similarity 33.6%; Pred. No. 8e-33;
Matches 118; Conservative 53; Mismatches 105; Indels 75; Gaps 8;
Qy 42 YTLQYIGEGAYGMVSAIDHVRKTRVAIKI-SPFEHQTYCQRTLEIQLRFRHENV 100
Db 25 FEKLEQIGEGTYGVYNAKKTGEIVALKKIRMDNEREGFPITAREIKILKKLHENV 84
Qy 101 IGIRDIILRASTLEMRD-----VYIVQDMETDLYKLL--KSQQLSNDHICYFL 147
Db 85 IKLKEIVTSPGPEKDEQGRPEGKYGKIYVFEYMDHDLTGLADRFPMRFSVPQIKCYM 144
Qy 148 YQILRGLKYTHSANVLRDLKPSNLLINTTCDLKIADFGlareVLTSPYVAT 207
Db 145 RQLITGLHYCHINQVLRDILKNSLLIDNEGNLADFGLAR--SFNDENANLITRVIT 202
Qy 208 RWRAPETIMNSKGYTKSIDISVGCILAEMLSNRPIFPKGHYLDQLNHIL----- 258
Db 203 LWYRPELLGATKYGPAVDWMSVGCIFAEILHGKPIFPKGDEPEQLNKIFELCGAPDEI 262
Qy 259 -----ALDLLDRMLTFNPNKRITVEEALAH 284
Db 263 NWPGVSKIPWYNFKPTRPMKRLREVPFRHFRHLELLEMLTLDPSQISAKDALD-- 320
Qy 285 YLEQYV-----DPTDEPVAEEPTFAMELDDLKRELKELIFQETARFQ 328
Db 321 --AEYFWADPLPCDPKSLPKYESSHEFQTK-----KKRQQORQHEETARQ 364

Search completed: February 6, 2006, 15:41:22
Job time : 11.1369 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:19:31 ; Search time 286.062 Seconds
(without alignments)
551.408 Million cell updates/sec

US-10-623-108-4

Title:
Perfect score: 1879

Sequence: 1 MAAAAAGGGGGPRRTGV.....LKELIFQTARFQGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	100.0	359	9	ADW12903 Human SMA
2	1879	100.0	359	9	ADW12907 Human SMA
3	1873	99.7	359	6	ABU54596 Human NOV
4	1859	98.9	379	2	AAW15506 Mitogen a
5	1859	98.9	379	4	AAM38754 Human pol
6	1859	98.9	379	6	ABR43898 Human ERK
7	1859	98.9	379	7	ADP45050 Human kin
8	1859	98.9	379	7	ADH59632 Erk1 prot
9	1859	98.9	379	8	ADO22522 Biochemic
10	1859	98.9	379	8	ADQ37846 Human pho
11	1859	98.9	379	9	ADV42070 Human ext
12	1859	98.9	379	9	ADW12905 Human SMA
13	1859	98.9	379	9	ADZ65036 Mitogen a
14	1859	98.9	403	4	AAM40540 Human pol
15	1853	98.6	379	2	AAI42413 Extracell
16	1853	98.6	379	4	AAG67618 Amino aci
17	1853	98.6	379	4	AAG67439 Amino aci
18	1853	98.6	379	5	ABG31847 Human ext
19	1853	98.6	379	6	ABU89742 protein d
20	1853	98.6	379	7	ABR62625 Human pro
21	1853	98.6	379	8	ADO24425 Human PRO
22	1853	98.6	379	9	ADZ10062 Human bre
23	1853	98.6	631	2	AAW85006 Erk1-gree
24	1853	98.6	631	3	AAI70778 EGFP-Erk1

25	1787	95.1	624	2	AAW85015	Aaw85015 Erk1-gree
26	1777	94.6	375	6	ABP96053	Abp96053 Human pro
27	1735	92.3	356	2	AAW48013	Aam48013 Extracell
28	1734	92.3	355	7	ADC64458	Adc64458 Human ext
29	1734	92.3	355	8	ADG98194	Adg98194 Human ext
30	1723	91.7	365	2	AAE20103	Aar20103 ERK1. 8/2
31	1614.5	85.9	340	8	ABM82549	Abm82549 Human dia
32	1614.5	85.9	340	8	ABM82547	Abm82547 Human dia
33	1599	85.1	335	8	ADH48364	Adh48364 Human KPP
34	1599	85.1	335	8	ABM82550	Abm82550 Human dia
35	1599	85.1	335	8	ABM82548	Abm82548 Human dia
36	1599	85.1	335	9	ADW12901	Adw12901 Human SMA
37	1548.5	82.4	380	7	ADC64459	Adc64459 Human ext
38	1548.5	82.4	380	8	ADG98195	Adg98195 Human ext
39	1538.5	81.9	360	5	ABG31848	Abg31848 Human cer
40	1538.5	81.9	360	6	ABR92109	AbR92109 Human cer
41	1538.5	81.9	360	6	ABR43899	ABr43899 Human ERK
42	1538.5	81.9	360	7	ADP45051	Adh45051 Human kin
43	1538.5	81.9	360	7	ADH59634	Adh59634 Erk2 prot
44	1538.5	81.9	360	8	ADJ66608	Adj66608 ERK prote
45	1538.5	81.9	360	8	ADP56268	Adp56268 Human PRO

ALIGNMENTS

RESULT 1

ADW12903
ID ADW12903 standard; protein; 359 AA.

XX AC ADW12903;

XX XX 07-APR-2005 (first entry)

DT XX Human SMAPK3V2 variant protein.

DE XX DNA purification; diagnosis; cancer; mitogen-activated protein kinase; enzyme.

KW XX Homo sapiens.

OS US2005013817-A1.

PN 20-JAN-2005.

XX PF 18-JUL-2003; 2003US-00623108.

XX PR 18-JUL-2003; 2003US-00623108.

XX PA (DAIK/) DAI K.

XX PI Dai K;

XX DR WPI; 2005-080923/09.

XX DR N-PSDB; ADW12902.

XX PT New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene in a mammal.

XX PS Claim 1; SEQ ID NO 4; 55pp; English.

XX CC The invention relates to novel isolated polypeptide (I) comprising an amino acid sequence selected from sequences comprising 335 or 359 amino acids (ADW12901 or ADW12903) or its fragments. The polypeptides and polynucleotides are useful for diagnosing diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene in a mammal. The fragments of the polypeptides and polynucleotides can also be used as primers or probes. This sequence corresponds to the SMAPK3V1 variant protein.

XX SQ Sequence 359 AA;

Query Match	100.0%;	Score 1879;	DB 9;	Length 359;
Best Local Similarity	100.0%;	Pred. No. 8.2e-187;		
Matches 359;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAAAAAGGGGEPRTTEGGVPGVPGEVEMVKGPFDVGPRTYQIQVIGEGAYGMWSAY	60	
Db	1	MAAAAAAGGGGEPRTTEGGVPGVPGEVEMVKGPFDVGPRTYQIQVIGEGAYGMWSAY	60	
Qy	61	DHVKTRVAIKKISPFESHQTYCQRTLREIQILLFRHENVIGIRDIILRASTLEAMRDVYI	120	
Db	61	DHVKTRVAIKKISPFESHQTYCQRTLREIQILLFRHENVIGIRDIILRASTLEAMRDVYI	120	
Qy	121	VQDLMETDLYKLLKSQLSNDHCYFYQIILRGULKYIHSANVLHRDLKPSNLLINTTCDL	180	
Db	121	VQDLMETDLYKLLKSQLSNDHCYFYQIILRGULKYIHSANVLHRDLKPSNLLINTTCDL	180	
Qy	181	KICDFGLARADPEHDHTGFLTEVATRVTRWRAPEIMLNSKGYTKSIDIWSVGCILAEMLS	240	
Db	181	KICDFGLARADPEHDHTGFLTEVATRVTRWRAPEIMLNSKGYTKSIDIWSVGCILAEMLS	240	
Qy	241	NRPIFPCKGHYLDQLNHILGILGSPSQBEDLNCIINWKARNYIQLSPSKTKVAAWKLFPKSD	300	
Db	241	NRPIFPCKGHYLDQLNHILGILGSPSQBEDLNCIINWKARNYIQLSPSKTKVAAWKLFPKSD	300	
Qy	301	SKALDLDRMLTFNPNKRITVAESPFTFAMELDDLPERLKELIIFOETARFPQGVLEAP	359	
Db	301	SKALDLDRMLTFNPNKRITVAESPFTFAMELDDLPERLKELIIFOETARFPQGVLEAP	359	

RESULT 2	
ADW12907	
ID	ADW12907 standard; protein; 359 AA.
XX	
AC	ADW12907;
XX	
DT	07-APR-2005 (first entry)
XX	
DE	Human SMAPK3V4 variant protein.
XX	
XX	DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW	enzyme.
KW	

[illegible]

DT 11-AUG-2003 (first entry)
XX Human ERK1 protein.
XX DARPP-32; psychotic disorder; intracellular signaling protein;
KW dopamine- and cAMP-regulated phosphoprotein; ERK1; ERK2; CREB;
KW extracellular signal-regulated protein kinase; phosphorylation;
KW cAMP-response element binding protein; schizophrenia; human.
XX Homo sapiens.
XX WO2003021225-A2.
PN 13-MAR-2003.
XX 03-SEP-2002; 2002WO-US027802.
XX 31-AUG-2001; 2001US-0316338P.
PR (UYRQ) UNIV ROCKEFELLER.
PA (KARO-) KAROLINSKA INST.
XX Greengard P, Pisone G;
PI WPI; 2003-300912/29.
XX Identifying agent to be tested for ability to treat psychotic disorder,
XX by contacting cells/tissues with candidate drug, determining
PT phosphorylation levels of intracellular signaling proteins DARPP-32,
PT ERK1, ERK2, CREB.
XX Disclosure; Page 77-78; 79pp; English.
XX The invention relates to identifying an agent to be tested for ability to
CC treat psychotic disorder in patient. The method involves contacting cells
CC or tissues with a candidate drug, and determining levels of
CC phosphorylation of intracellular signaling proteins DARPP-32 (dopamine-
CC and cAMP-regulated phosphoprotein), ERK1 and ERK2 (extracellular signal-
CC regulated protein kinases 1 and 2), and CREB (cAMP-response element
CC binding protein). The method is useful for identifying an agent to be
CC tested for an ability to treat a psychotic disorder such as schizophrenia
CC in a patient in need of such treatment. The present sequence represents a
CC human ERK1 protein
XX Sequence 379 AA;
Query Match 98.9%; Score 1859; DB 6; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAAGGGGGGPRTEGVPGEVEMVKQPPDVGPRYTQLOVIGGAYGMVSSAY 60
DB 1 MAAAAAGGGGGGPRTEGVPGEVEMVKQPPDVGPRYTQLOVIGGAYGMVSSAY 60
QY 61 DHVKTTRVAIKKISPFPHQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTTRVAIKKISPFPHQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEML 240
DB 181 KICDFGLARADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEML 240
QY 241 NRPIFPKGHYLDQLNHLIGLSPSQEDLNCIINKARNYLQSLPSKTKVAKLPPKSD 300
DB 241 NRPIFPKGHYLDQLNHLIGLSPSQEDLNCIINKARNYLQSLPSKTKVAKLPPKSD 300
QY 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKRL 340
DB 301 SKALDLDLRLMTFNPKNRITVEALAHPLYEQYYDPTDEPVAEPPTFAMELDDLPKRL 360

QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 7
ADP45050
ID ADP45050 standard; protein; 379 AA.
XX ADF45050;
XX 12-FEB-2004 (first entry)
DT Human kinase ERK1.
DE Human; protein kinase; enzyme; inhibitor; ERK1.
KW Homo sapiens.
XX WO2003081210-A2.
PN 02-OCT-2003.
XX 20-MAR-2003; 2003WO-US008725.
XX 21-MAR-2002; 2002US-0366892P.
XX (SUNE-) SUNESIS PHARM INC.
PA Prescott JC, Braisted A;
PI WPI; 2003-865136/80.
XX Identifying ligand binding to inactive conformation of target protein
XX kinase (T) comprises contacting the conformation modified (T) which
PT contains reactive group at binding site, with ligands and detecting
PT kinase-ligand conjugate formation.
XX Disclosure; SEQ ID NO 19; 260pp; English.
XX The present invention relates to a method for identifying a ligand (L),
CC which binds to an inactive conformation of target protein kinase (T). The
CC method involves contacting inactive conformation of (T), which contains
CC or is modified to contain a reactive group at or near a binding site of
CC interest, with one or more ligand candidates capable of covalently
CC bonding to the reactive group thus forming a kinase-(L) conjugate (C).
CC The method is useful for identifying protein kinase inhibitors that
CC preferentially bind to inactive conformation of a target protein kinase.
CC The present sequence is a protein kinase which may be modified via an
CC amino acid substitution, for use in the method of the invention.
XX Sequence 379 AA;
Query Match 98.9%; Score 1859; DB 7; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAAGGGGGGPRTEGVPGEVEMVKQPPDVGPRYTQLOVIGGAYGMVSSAY 60
DB 1 MAAAAAGGGGGGPRTEGVPGEVEMVKQPPDVGPRYTQLOVIGGAYGMVSSAY 60
QY 61 DHVKTTRVAIKKISPFPHQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTTRVAIKKISPFPHQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEML 240
DB 181 KICDFGLARADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEML 240

QY 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
DB 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPPFTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNKRITVEEALAHFVLEQYDDTDFVAAEPPFTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 8
ADH59632
ID ADH59632 standard; protein; 379 AA.
XX
AC ADH59632;
XX
XX 25-MAR-2004 (first entry)
XX
DE Erk1 protein.
XX
XX cardiac disease; Raf-1; MEK1; Cardiant; telethonin; &agr-actinin;
KW hypertrophy; MEK2; Erk1; Erk2.
XX
XX Homo sapiens.
XX
XX WO2003025205-A2.
XX
XX 27-MAR-2003.
XX
XX 18-SEP-2002; 2002WO-EP010489.
XX
XX 19-SEP-2001; 2001US-0323566P.
PR
XX 24-SEP-2001; 2001US-0324625P.
XX
XX (MEDI-) MEDIGENE AG.
PA
XX Nave B, Roenicke V, Leclair S, Funk M, Reuner B, Brinkmann K;
PI Henkel T;
XX
XX WPI: 2003-371821/35.
DR N-PSDB; ADH59631.
DR

XX Identifying and/or obtaining a compound useful for preventing or treating
PT cardiac diseases, particularly congestive heart failure, comprises
PT quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
PT on the compound.
XX
XX Claim 33; SEQ ID NO 8; 41pp; English.
XX
XX The present invention relates to identifying and obtaining a compound
CC useful in the prevention or treatment of cardiac diseases, comprising
CC quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
CC on the compound. The method is useful for identifying and/or obtaining
CC compounds that may be used in the prevention or treatment of cardiac
CC diseases, particularly congestive heart failure. The compound or protein
CC is also used for the preparation of a pharmaceutical composition for
CC prevention or treatment of a disease related to hypertrophy or impaired
CC or increased activation of telethonin (T-Cap), &agr;-actinin, MHC, actin,
CC titin, myomesin, nebulin, tropomyosin, troponin, Erk1/2 and/or MLCK.
CC The protein or antibody which specifically recognizes the
CC activated/phosphorylated form of the above polypeptide, is used for the
CC preparation of a composition for diagnosing a disease or a predisposition
CC for a disease related to hypertrophy or related to impaired or increased
CC activation of Raf-1, MEK1/2 and/or Erk1/2. The present sequence
CC represents Erk1.
XX
XX Sequence 379 AA;
SQ
Query Match 98.9%; Score 1859; DB 7; Length 379;

Best Local Similarity 94.7%; Pred. NO. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAOQGGGGEPRTTEGVGPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
DB 1 MAAAAOQGGGGEPRTTEGVGPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTTCORTLREIQILLRPHENVIGIRDIRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQTTCORTLREIQILLRPHENVIGIRDIRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICVFYQIILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICVFYQIILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSKIDISVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSKIDISVGCILAEMLS 240
QY 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
DB 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPPFTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNKRITVEEALAHFVLEQYDDTDFVAAEPPFTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379
RESULT 9
AD022522
ID AD022522 standard; protein; 379 AA.
XX
AC AD022522;
XX
DT 12-AUG-2004 (first entry)
XX
DE Biochemical pathway-related human ERK protein SeqID8.
XX
KW biochemical pathway; mammalian glioma tumour;
KW epidermal growth factor receptor; EGFR; mTOR; polypeptide inhibitor;
KW cancer; ERK; human.
XX
OS Homo sapiens.
XX
XX WO2004044218-A2.
XX
PD 27-MAY-2004.
XX
PF 05-NOV-2003; 2003WO-US0351115.
XX
PR 05-NOV-2002; 2002US-0423777P.
XX
PA (REGC) UNIV CALIFORNIA.
XX (CELL-) CELL SIGNALING TECHNOLOGY INC.
XX
PI Mischel PS, Sawyers CL, Smith BL, Crosby K;
XX WPI; 2004-411736/38.
DR
XX Use of biochemical pathways associated with glioblastoma for, e.g.
PT identifying a mammalian glioma tumor that is likely to respond to an
PT epidermal growth factor receptor polypeptide inhibitor or an mTOR
PT polypeptide inhibitor.
XX
XX Claim 1; SEQ ID NO 8; 66pp; English.
PS
XX This invention relates to a novel use of biochemical pathways for
CC identifying a mammalian glioma tumor that is likely to respond to an
CC epidermal growth factor receptor (EGFR) polypeptide inhibitor or an mTOR

CC polypeptide inhibitor, or identifying a mammalian glioma tumour that does
CC not express or expresses a PTEN polypeptide and which is likely to
CC respond or not likely to respond to an inhibitor of mTOR polypeptide
CC activity or inhibitor of EGFR polypeptide activity, respectively. The
CC biochemical pathways are, in particular, dysregulated in pathologies such
CC as cancer. The present sequence is that of the human ERK protein which
CC may be used in the method of the invention.
XX
SQ Sequence 379 AA;

Query Match 98.9%; Score 1859; DB 8; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFQDVGPRTYQIQVIGEGYGMVSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFQDVGPRTYQIQVIGEGYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPPFHHQYTCQRTLEIQLLRFRHENVIGIRDLASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPPFHHQYTCQRTLEIQLLRFRHENVIGIRDLASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSKQSLNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSKQSLNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
DB 241 NRPIFGKHVLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
QY 301 SKALDLLDRMLTFNPNKRIT-----VASEPFTFAMELDDLPKRL 340
DB 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDFVASEPFTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 10
ADQ37846
ID ADQ37846 standard; protein; 379 AA.
XX
AC ADQ37846;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human phosphorylated ERK protein, target for HER-2 therapy SeqID 3.
XX
KW human; HER-2 directed therapy; tumour; insulin growth factor receptor;
KW IGFR; epidermal growth factor receptor; EGFR; S6 ribosomal protein; AKT;
KW NDF; ERK; cancer therapy; predictive biomarker; HER-2/neu.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 202 /note= "Optionally a phosphorylated serine (if threonine
FT at 204 is unmodified)"
FT Modified-site 204 /note= "Optionally a phosphorylated threonine (if serine
FT at 202 is unmodified)"
XX
PN WO2004053497-A2.
XX
PD 24-JUN-2004.
XX
PF 11-DEC-2003; 2003WO-US039770.
XX

PR 11-DEC-2002; 2002US-0432942P.
XX
PA (VENT-) VENTANA MEDICAL SYSTEMS INC.
PA (CELL-) CELL SIGNALING TECHNOLOGY INC.
XX
PI Bacus SS, Smith BL;
XX
DR WPI; 2004-507069/48.
XX
PT Identifying a mammalian tumor, useful for predicting a response to HER2-
PT directed therapy, comprises assaying a sample to detect a pattern of
PT expression, phosphorylation or both, of one or more polypeptides.
XX
PS Claim 35; SEQ ID NO 3; 49pp; English.
XX
CC This invention relates to a novel method for determining or predicting
CC the response of a patient to HER-2 directed therapy. Specifically, it
CC refers to analysing a mammalian tumour in order to detect a pattern of
CC expression and/or phosphorylation of a protein taken from the group
CC including insulin growth factor receptor (IGFR) polypeptide, epidermal
CC growth factor receptor (EGFR), phosphorylated NDF or phosphorylated ERK protein. The
CC present invention describes characterising a mammalian tumour's
CC responsiveness to an HER-2 therapy and hence an individual's response to
CC this cancer therapy, using an immunologically specific antibody directed
CC against one of the aforementioned proteins. Furthermore, it provides
CC predictive biomarkers that can be used to assess the efficacy of
CC therapeutic agents targeted to HER-2/neu. This polypeptide sequence is
CC the human phosphorylated ERK protein of the invention.
XX
SQ Sequence 379 AA;

Query Match 98.9%; Score 1859; DB 8; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFQDVGPRTYQIQVIGEGYGMVSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFQDVGPRTYQIQVIGEGYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPPFHHQYTCQRTLEIQLLRFRHENVIGIRDLASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPPFHHQYTCQRTLEIQLLRFRHENVIGIRDLASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSKQSLNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSKQSLNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
DB 241 NRPIFGKHVLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
QY 301 SKALDLLDRMLTFNPNKRIT-----VASEPFTFAMELDDLPKRL 340
DB 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDFVASEPFTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 11
ADV42070
ID ADV42070 standard; protein; 379 AA.
XX
AC ADV42070;
XX
DT 24-FEB-2005 (first entry)
XX

QY 1 MAAAAAQQGGGGGPRRTTEGVGPGVEVMVKGPFDVGPRTYQLOQYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGGPRRTTEGVGPGVEVMVKGPFDVGPRTYQLOQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
QY 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLQSLPSKTKVAAWAKLPKSD 300
DB 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLQSLPSKTKVAAWAKLPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYVPTDEPVAEPTTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379
RESULT 13
ID ADZ65036 standard; protein; 379 AA.
AC ADZ65036;
XX ADZ65036;
DT 14-JUL-2005 (first entry)
XX Mitogen activated protein kinase, MAPK-3, SEQ ID 2.
XX MAP kinase; Mitogen-activated protein kinase; enzyme; cancer.
XX Homo sapiens.
XX US2005095657-A1.
XX 05-MAY-2005.
XX 21-SEP-2004; 2004US-00945684.
XX 11-OCT-2002; 2002US-0418038P.
XX 14-OCT-2003; 2003WO-US032248.
XX (ARBI/) ARBISER J L.
PA (COHE/) COHEN C.
XX Arbiser JL, Cohen C;
XX WPI; 2005-354474/36.
DR REFSEQ; XP_055766.3.
XX Detecting phosphorylated mitogen activated protein kinase (P-MAPK), by
PT contacting sample with antibody having affinity for P-MAPK, detecting
PT antibody/P-MAPK complex to detect P-MAPK, and expression of P-MAPK
PT indicates cancer.
XX Claim 3; SEQ ID NO 2; 23pp; English.
XX The present invention relates to a method (M1) for detecting
CC Phosphorylated Mitogen Activated Protein Kinase (P-MAPK; ADZ65035-
CC ADZ65042), by providing sample, contacting the sample with at least one
CC antibody having affinity for the phosphorylated portion of P-MAPK, and
CC detecting antibody/P-MAPK complex, in which the presence of antibody/P-
CC MAPK complex indicates that P-MAPK is present in sample, and where

CC constitutive expression of P-MAPK is indicative of cancer.
XX
SQ Sequence 379 AA;
Query Match 98.9%; Score 1859; DB 9; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAAQQGGGGGPRRTTEGVGPGVEVMVKGPFDVGPRTYQLOQYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGGPRRTTEGVGPGVEVMVKGPFDVGPRTYQLOQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
QY 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLQSLPSKTKVAAWAKLPKSD 300
DB 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLQSLPSKTKVAAWAKLPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYVPTDEPVAEPTTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379
RESULT 14
AAW40540
ID AAW40540 standard; protein; 403 AA.
XX AAW40540;
XX AAW40540;
DT 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 5471.
XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
PA

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI59696.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PT
XX Example 2; SEQ ID NO 5471; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAW38442-AAW42213) with nootropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX Sequence 403 AA;
XX
Query Match 98.9%; Score 1859; DB 4; Length 403;
Best Local Similarity 94.7%; Pred. No. 1.2e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAQAQGGGGPRRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
DB 25 MAAAQAQGGGGPRRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 84
QY 61 DHVTRTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
DB 85 DHVTRTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDIRASTLEAMRDVYI 144
QY 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 145 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 204
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 205 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 264
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPDKTKVAKLPFKSD 300
DB 265 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPDKTKVAKLPFKSD 324
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPTEFAMELDDLPKRL 340
DB 325 SKALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDBPVAEPPTEFAMELDDLPKRL 384
QY 341 KELIFOETARFQGVLEAP 359
DB 385 KELIFOETARFQGVLEAP 403
RESULT 15
AAV42413
ID AAV42413 standard; protein; 379 AA.
XX
AC AAV42413;
XX
DT 02-DEC-1999 (first entry)
XX
DE Extracellular signal Regulated Kinase (ERK)1 mutant.

XX
KW mitogen activated protein; MAP kinase; apoptosis; cancer; inflammation;
KW intracellular signal transduction pathway; inhibitor; wildtype; ERK;
KW extracellular signal regulated kinase; pyridinyl-imidazole.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 122
FT /note= "Glu can be substituted by an amino acid with a
FT small side chain, for example alanine or threonine"
XX
XX W09942592-A1.
XX
XX 26-AUG-1999.
XX
XX 16-FEB-1999; 99WO-US003181.
XX
XX 18-FEB-1998; 98US-00025580.
XX (VERT-) VERTEX PHARM INC.
XX Su MS, Fox E, Wilson KP, Germann UA;
XX WPI; 1999-540310/45.
XX Method of designing Ser/Thr or Tyr kinase inhibitor useful for treating,
PT example breast cancer, restenosis, asthma or hypertension.
XX
XX Example 6; Page 50-51; 71pp; English.
XX
XX This is the amino acid sequence of ERK1. Substitution of glutamine with
CC an amino acid with a small side chain, such as alanine and threonine will
CC create a mutant kinase which can theoretically bind to a pyridinyl-
CC imidazole inhibitor of p38. (MAP) 1 kinase. MAP 1 kinases mediate
CC intracellular signal transduction pathways and so have a role in many
CC diverse human diseases. For example, kinases have been implicated in cell
CC entry into apoptosis, cancer, Alzheimer's disease, angiotensin II and
CC hematopoietic cytokine receptor signal transduction, oncoprotein
CC signalling and mitosis, inflammation and infection, etc. Members of the
CC MAP kinase family share sequence similarity and conserved structural
CC domains, and include the extracellular-signal regulated kinases (ERKs),
CC Jun N-terminal kinases (JNKs) and p38 kinases. The invention relates to
CC methods for designing inhibitors of serine/threonine kinases,
CC particularly MAP kinases, and tyrosine kinases through the use of ATP-
CC binding site mutants of these kinases. The methods of this invention take
CC advantage of the fact that the mutant kinases are capable of binding
CC inhibitory compounds of other kinases with greater affinity than the
CC corresponding wild-type kinase
XX
XX Sequence 379 AA;
XX
Query Match 98.6%; Score 1853; DB 2; Length 379;
Best Local Similarity 94.5%; Pred. No. 4.6e-184;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;
QY 1 MAAAQAQGGGGPRRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
DB 1 MAAAQAQGGGGPRRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
QY 61 DHVTRTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
DB 61 DHVTRTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
QY 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPDKTKVAKLPFKSD 300

Db	241	NRPIFGKHYLDQNHILGILGSPQEDLNCIINMKARNYLQSLPSKTKVAAKLPKSD	300
Qy	301	SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKERL	340
Db	301	SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVAEPEPTFAMELDDLPKERL	360
Qy	341	KELIFOETARFPQGVLEAP	359
Db	361	KELIFOETARFPQGVLEAP	379

Search completed: February 6, 2006, 15:29:02
Job time : 287.562 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:29:28 ; Search time 26.8991 Seconds
(without alignments)
1284.124 Million cell updates/sec

Title: US-10-623-108-4
Perfect score: 1879
Sequence: 1 MAAAAAGGGGGPRRTGV.....LKELIFQETARFGVLEAP 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1853	98.6	379	1 A48082	MAP kinase 3 (EC 2
2	1784	94.9	380	2 JCI451	Ca2+/calmodulin-de
3	1783	94.9	377	2 S28184	Ca2+/calmodulin-de
4	1767	94.0	369	2 A60041	Ca2+/calmodulin-de
5	1550.5	82.5	392	2 JW0052	extracellular sign
6	1538.5	81.9	360	1 JQ1400	MAP kinase 1 (EC 2
7	1538.5	81.9	360	2 S25011	protein kinase ERK
8	1528.5	81.3	358	2 S16444	mitogen-activated
9	1528.5	81.3	358	2 A40033	protein kinase (EC
10	1516	80.7	361	2 A39754	mitogen-activated
11	1508	80.3	369	2 JW0053	extracellular sign
12	1364.5	72.6	376	2 A46036	extracellular sign
13	1317	70.1	376	2 A36978	MAP kinase mpk-1 (
14	1317	70.1	444	2 A36977	MAP kinase sur-1 (
15	932	49.6	356	2 T51944	pathogenicity MAP
16	918.5	48.9	361	2 T51943	mitogen-activated
17	884	47.0	372	2 S15663	protein kinase (EC
18	865	46.0	415	2 A56042	mitogen-activated
19	859.5	45.7	362	2 A47211	protein kinase ERK
20	837.5	44.6	387	2 S48123	mitogen-activated
21	827	44.0	394	2 S33635	mitogen-activated
22	826.5	44.0	371	2 S51320	mitogen-activated
23	822	43.7	376	2 F96619	protein T30E16.13
24	822	43.7	395	2 S40472	mitogen-activated
25	817	43.5	393	2 S51321	mitogen-activated
26	814	43.3	372	2 S39559	mitogen-activated
27	814	43.3	384	2 S52989	mitogen-activated
28	813	43.3	376	2 S40470	mitogen-activated
29	807.5	43.0	815	2 B56708	extracellular sign

30	806.5	42.9	369	2 S56638	mitogen-activated
31	806.5	42.9	370	2 T47504	mitogen-activated
32	806.5	42.9	370	2 S40469	mitogen-activated
33	803	42.7	368	2 A33297	probable serine/th
34	801.5	42.7	363	2 C86214	hypothetical prote
35	801	42.6	353	2 S28548	protein kinase (EC
36	800	42.6	393	2 T47803	mitogen-activated
37	797	42.4	370	2 F86236	protein FLN23.9 [
38	793.5	42.2	376	2 S40471	mitogen-activated
39	791	42.1	375	2 T03971	mitogen-activated
40	789.5	42.0	371	2 S60121	mitogen-activated
41	786	41.8	368	2 S40473	mitogen-activated
42	785.5	41.8	371	2 T09622	protein kinase MMK
43	783	41.7	368	2 B84561	MAP kinase (ATMPK7
44	782.5	41.6	371	2 T14915	mitogen-activated
45	781.5	41.6	369	2 C86146	hypothetical prote

ALIGNMENTS

RESULT 1

A48082
MAP kinase 3 (EC 2.7.1.1) - human
N;Alternate names: extracellular signal-regulated kinase 1 (ERK1); mitogen-activated protein kinase (EC 2.7.1.37)
C;Species: Homo sapiens (man)
C;Date: 04-Sep-1998 #sequence, revision 04-Sep-1998 #text_change 09-Jul-2004
C;Accession: A48082; PQ0270; S23428; S15519; S21579
R;Character: D.L.; Mordret, G.; Harder, K.W.; Jirik, F.; Pelech, S.L.
Mol. Cell. Biol. 13, 4679-4690, 1993
A;Title: Molecular cloning, expression, and characterization of the human mitogen-activated protein kinase 3 (p44^{ERK1})
A;Reference number: JQ1400; MUID:93330262; PMID:7687743
A;Accession: A48082
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-379 <CHA>
A;Cross-references: UNIPROT:P27361; UNIPARC:UPI0000035BE2; EMBL:X60188; NID:g31220; PIDN:CAA77754.1; PID:9233882
A;Experimental source: hepatoma cell line HEP G2
A;Note: authors translated the codon AGC for residue 174 as Ile
R;Owaki, H.; Makar, R.; Boulton, T.G.; Cobb, M.H.; Geppert, T.D.
Biochem. Biophys. Res. Commun. 182, 1416-1422, 1992
A;Title: Extracellular signal-regulated kinases in T cells: characterization of human ERK1
A;Reference number: JQ1400; MUID:92171961; PMID:1540184
A;Accession: PQ0270
A;Molecule type: mRNA
A;Residues: 14-173, 'I', 175-379 <OWA>
A;Cross-references: UNIPARC:UPI000016AB99; GB:M84490; NID:g186695; PIDN:AAA36142.1; PID:9233882
A;Experimental source: cell line CEM
R;Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.
FEBS Lett. 304, 170-178, 1992
A;Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.
A;Reference number: S23426; MUID:92316223; PMID:1319925
A;Accession: S23428
A;Molecule type: mRNA
A;Residues: 25-173, 'I', 175-379 <GON>
A;Cross-references: UNIPARC:UPI000016A2B4; EMBL:Z11696; NID:g233882; PIDN:CAA77754.1; PID:9233882
C;Comment: This enzyme is activated by MAP kinase kinase (see PIR:45100 and PIR:A46723).
C;Genetics:
A;Gene: GDB:PRKM3; ERK1
A;Cross-references: GDB:135679; OMIM:601795
A;Map position: 16pter-16qter
C;Complex: monomer
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate
A;Pathway: MAP kinase cascade
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; monomer; phosphoprotein; phosphotransferase; serine/threonine-specific kinase
F;40-330/Domain: protein kinase homology <KIN>
F;48-56/Region: protein kinase ATP-binding motif
F;202/Binding site: phosphate (Thr) (covalent) (by MAP kinase kinase) #status predicted
F;204/Binding site: phosphate (Tyr) (covalent) (by MAP kinase kinase) #status predicted

Query Match	98.6%;	Score 1853;	DB 1;	Length 379;
Best Local Similarity	94.5%;	Pred. No. 1.1e-77;		
Matches 359;	Conservative 0;	Mismatches 1;	Indels 20;	Gaps 1;

QY	1	MAAAAQGGGGGPRRTGEGVGPVGEVEMVKQPPDFVGPRTYQLQYIGEGAYGMVSSAY	60
DB	1	MAAAAQGGGGGPRRTGEGVGPVGEVEMVKQPPDFVGPRTYQLQYIGEGAYGMVSSAY	60
QY	61	DHVKTRVAIKKISPPEHQTQYQRTIREIQILRFHENVIGIRDILRASTLEAMRDVYI	120
DB	61	DHVKTRVAIKKISPPEHQTQYQRTIREIQILRFHENVIGIRDILRASTLEAMRDVYI	120
QY	121	VDLMTDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSANVLHRDLKPSNLLINTTCDL	180
DB	121	VDLMTDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSANVLHRDLKPSNLLSNTTCDL	180
QY	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS	240
DB	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS	240
QY	241	NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSD	300
DB	241	NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSD	300
QY	301	SKALDLLDRMLTFNPKNRIT-----VAEPTTFAMELDDLPERLK	340
DB	301	SKALDLLDRMLTFNPKNRITVEEALAHPLYEQYVDTDEPVAEPTTFAMELDDLPERLK	360
QY	341	KELIFQETARFQGVLEAP	359
DB	361	KELIFQETARFQGVLEAP	379

RESULT 2

JC1451

Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) - rat

N/Alternate names: ERK1-MAP kinase; extracellular signal-regulated kinase

C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004

C/Accession: JC1451; A35061; A37140; A40466; S24947

R/Marquardt, B.; Stabel, S.

Gene 120, 297-299, 1992

A/Title: Sequence of a rat cDNA encoding the ERK1-MAP kinase.

A/Reference number: JC1451; MUID:93013050; PMID:1327976

A/Accession: JC1451

A/Molecule type: mRNA

A/Residues: 1-380 <WAR>

A/Cross-references: UNIPROT:P21708; UNIPARC:UPI000012P174; EMBL:X65198; NID:G56626; PIDN:G56626

A/Experimental source: brain

R/Boulton, T.G.; Yancopoulos, G.D.; Gregory, J.S.; Slaughter, C.; Moornaw, C.; Hsu, J.; C

Science 249, 64-67, 1990

A/Title: An insulin-stimulated protein kinase similar to yeast kinases involved in cell

A/Reference number: A35061; MUID:90312137; PMID:2164259

A/Accession: A35061

A/Molecule type: mRNA

A/Residues: 14-380 <BOU>

A/Cross-references: UNIPARC:UPI0000145072; GB:M38194; NID:G204051; PIDN:AAA41123.1; PID:G204051

R/Boulton, T.G.; Gregory, J.S.; Cobb, M.H.

Biochemistry 30, 278-286, 1991

A/Title: Purification and properties of extracellular signal-regulated kinase 1, an insu

A/Reference number: A37140; MUID:91105092; PMID:1846291

A/Accession: A37140

A/Status: preliminary

A/Molecule type: protein

A/Residues: 43-64;167-178, 'X',180-183, 'X',185 <BO2>

A/Cross-references: UNIPARC:UPI000017559F; UNIPARC:UPI0000175590

R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.

DNA Cell Biol. 10, 505-514, 1991

A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse a

A/Reference number: A40466; MUID:91369479; PMID:1716439

A/Accession: A40466

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 14-94, 'R', 96-380 <DEM>					
A/Cross-references: UNIPARC:UPI00001707CE; GB:U12008; GB:S59509; NID:G515498; PIDN:AAA201					
C/Superfamily: kinase-related transforming protein; protein kinase homology					
C/Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; serine/threonine					
F:41-331/Domain: protein kinase homology <KIN>					
F:49-57/Region: protein kinase ATP-binding motif					
Query Match		94.9%; Score 1784; DB 2; Length 380;			
Best Local Similarity		91.8%; Pred. No. 1.5e-74;			
Matches 347; Conservative		1; Mismatches 10; Indels 20; Gaps 1;			
QY	2	AAAAAQGGGGGPRRTGEGVGPVGEVEMVKQPPDFVGPRTYQLQYIGEGAYGMVSSAYD	61		
DB	3	AAAAAPGGGGGPRGTAGVVPVPGVEVVKQPPDFVGPRTYQLQYIGEGAYGMVSSAYD	62		
QY	62	HVKTRVAIKKISPPEHQTQYQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVVIV	121		
DB	63	HVKTRVAIKKISPPEHQTQYQRTIREIQILGFRHENVIGIRDILRAPTLEAMRDVVIV	122		
QY	122	QDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSANVLHRDLKPSNLLINTTCDLK	181		
DB	123	QDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSANVLHRDLKPSNLLINTTCDLK	182		
QY	182	ICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN	241		
DB	183	ICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN	242		
QY	242	RIPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSDS	301		
DB	243	RIPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSDS	302		
QY	302	KALDLLDRMLTFNPKNRIT-----VAEPTTFAMELDDLPERLK	341		
DB	303	KALDLLDRMLTFNPKNRITVEEALAHPLYEQYVDTDEPVAEPTTFDMELDDLPERLK	362		
QY	342	ELIFQETARFQGAPEAP	359		
DB	363	ELIFQETARFQGAPEAP	380		
RESULT 3					
S28184					
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) erk-1 - mouse (fragment)					
N/Alternate names: extracellular signal-regulated kinase 1					
C/Species: Mus musculus (house mouse)					
C/Date: 22-Nov-1993 #sequence revision 30-Jan-1998 #text_change 09-Jul-2004					
C/Accession: S28184; B40466; A41371					
R/Tanner, B.; Mueckler, M.					
Biochim. Biophys. Acta 1171, 319-320, 1993					
A/Title: Molecular cloning of a mouse extracellular signal regulated kinase (erk-1).					
A/Reference number: S28184; MUID:93144347; PMID:8424957					
A/Accession: S28184					
A/Molecule type: mRNA					
A/Residues: 1-377 <TAN>					
A/Cross-references: UNIPROT:Q63844; UNIPARC:UPI000017558C; EMBL:Z14249					
R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.					
DNA Cell Biol. 10, 505-514, 1991					
A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse					
A/Reference number: A40466; MUID:91369479; PMID:1716439					
A/Accession: B40466					
A/Molecule type: mRNA					
A/Residues: 74-377 <DEM>					
A/Cross-references: UNIPARC:UPI0000023599; GB:S59517					
R/Crews, C.M.; Alessandrini, A.A.; Erikson, R.L.					
Proc. Natl. Acad. Sci. U.S.A. 88, 8845-8849, 1991					
A/Title: Mouse Erk-1 gene product is a serine/threonine protein kinase that has the poten					
A/Reference number: A41371; MUID:92020947; PMID:1717989					
A/Accession: A41371					
A/Molecule type: mRNA					
A/Residues: 7-16 <CRE>					
A/Cross-references: UNIPARC:UPI0000170C57; GB:S58470; NID:G236372; PIDN:AAB19973.1; PID:G236372					
C/Superfamily: kinase-related transforming protein; protein kinase homology					
C/Keywords: ATP; calmodulin binding; phosphotransferase; protein kinase					

F;38-328/Domain: protein kinase homology <KIN>
F;46-54/Region: protein kinase ATP-binding motif

Query Match 94.8%; Score 1783; DB 2; Length 377;
Best Local Similarity 91.8%; Pred. No. 1.6e-74;
Matches 346; Conservative 1; Mismatches 10; Indels 20; Gaps 1;
QY 3 AAAAAQGGGGEPRRTGEGVGVGPEVEMVKGQFDVGPRTYQIQYIGEGAGMVSSAYDH 62
DB 1 AAAAAQGGGGEPRGTAGVDPVDPGEVEMVKGQFDVGPRTYQIQYIGEGAGMVSSAYDH 60
QY 63 VRKTRVAIKKISPEHQTYCQRTLEIQILRRHENHVGIRDLRASTLEAMRDVYIQ 122
DB 61 VRKTRVAIKKISPEHQTYCQRTLEIQILRRHENHVGIRDLRASTLEAMRDVYIQ 120
QY 123 DLMETDLYKLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI 182
DB 121 DLMETDLYKLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI 180
QY 183 CDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSIDISVGCILAEMLSNR 242
DB 181 CDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSIDISVGCILAEMLSNR 240
QY 243 PIPFGKHYLDQNLHILGILSPQSDQEDLNCIINNKARNYLOSLSKTKVAVAKLFPKSDSK 302
DB 241 PIPFGKHYLDQNLHILGILSPQSDQEDLNCIINNKARNYLOSLSKTKVAVAKLFPKSDSK 300
QY 303 ALDLDRLMLTFNPKRIT-----VAEPPFTFAMELDDLPKERLKE 342
DB 301 ALDLDRLMLTFNPKRITVEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERLKE 360
QY 343 LIQETARFQGVLEAP 359
DB 361 LIQETARFQGAPEGP 377

RESULT 4
A60041
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) ERK1 - Chinese hamster (fragment
N/Alternate names: extracellular signal-regulated kinase 1 (ERK1); MAP kinase; microtubu
C/Species: Cricetus griseus (Chinese hamster)
C/Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 09-Jul-2004
C/Accession: A60041
R/Meloche, S.; Pages, G.; Pouyssegur, J.
Mol. Biol. Cell 3, 63-71, 1992
A/Title: Functional expression and growth factor activation of an epitope-tagged p44 mit
A/Reference number: A60041; MUID:92199340; PMID:1372523
A/Accession: A60041
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-369 <MEL>
A/Cross-references: UNIPROT:Q7M0H9; UNIPARC:UPI0000175588
C/Comment: This serine/threonine kinase is activated in response to extracellular stimul
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; protein kinase
F;30-320/Domain: protein kinase homology <KIN>
F;38-46/Region: protein kinase ATP-binding motif

Query Match 94.0%; Score 1767; DB 2; Length 369;
Best Local Similarity 92.4%; Pred. No. 8.5e-74;
Matches 341; Conservative 1; Mismatches 7; Indels 20; Gaps 1;
QY 11 GGEPRRTGEGVGVGPEVEMVKGQFDVGPRTYQIQYIGEGAGMVSSAYDHVKRKTVAI 70
DB 1 GGEPRGAAGVGVGPEVEMVKGQFDVGPRTYQIQYIGEGAGMVSSAYDHVKRKTVAI 60
QY 71 KKISPFHQTYCQRTLEIQILRRHENHVGIRDLRASTLEAMRDVYIQDLMETDLY 130
DB 61 KKISPFHQTYCQRTLEIQILRRHENHVGIRDLRASTLEAMRDVYIQDLMETDLY 120
QY 131 KLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI CDFGLARI 190
DB 121 KLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI CDFGLARI 180

QY 191 ADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSIDISVGCILAEMLSNRPIPGKH 250
DB 181 ADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSIDISVGCILAEMLSNRPIPGKH 240
QY 251 LDQLNHILGILSPQSDQEDLNCIINNKARNYLOSLSKTKVAVAKLFPKSDSKALDLDLDM 310
DB 241 LDQLNHILGILSPQSDQEDLNCIINNKARNYLOSLSKTKVAVAKLFPKSDSKALDLDLDM 300
QY 311 LTFNPNKRIT-----VAEPPFTFAMELDDLPKERLKELIPOETAR 350
DB 301 LTFNPNKRITVEDALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERLKELIPOETAR 360
QY 351 FQPGVLEAP 359
DB 361 FQGAPEAP 369
RESULT 5
JW0052
extracellular signal-regulated kinase (EC 2.7.-.-) 1 - common carp
N/Alternate names: CERK1
C/Species: Cyprinus carpio (common carp)
C/Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C/Accession: JW0052
R/Hashimoto, H.; Yokoyama, Y.; Matsuo, Y.; Toyohara, H.; Kohno, M.; Sakaguchi, M.
J. Biochem. 123, 1031-1035, 1998
A/Title: Existence of two isoforms of extracellular signal-regulated kinase in fish.
A/Reference number: JW0052; MUID:98269030; PMID:9603989
A/Accession: JW0052
A/Molecule type: mRNA
A/Residues: 1-392 <HAS>
A/Cross-references: UNIPROT:Q7LZH4; UNIPARC:UPI0000175598
C/Comment: This enzyme is the key enzyme in the intracellular signal transduction pathway
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: phosphotransferase
F;54-344/Domain: protein kinase homology <KIN>

Query Match 82.5%; Score 1550.5; DB 2; Length 392;
Best Local Similarity 78.5%; Pred. No. 5.8e-64;
Matches 300; Conservative 20; Mismatches 29; Indels 33; Gaps 3;
QY 3 AAAAAQGGGGEPRRTGEGVGV-----PGEVEMVKGQFDVGPRTYQIQYIG 50
DB 6 SAAAAAGAGSSSSAAGAAVAVGANGAAGPKPG-LESVKGNQFVGPRTYDQYIG 64
QY 51 GAYGMVSSAYDHVKRKTVAIKKISPEHQTYCQRTLEIQILRRHENHVGIRDLRAS 110
DB 65 GAYGMVCSAFDNVKNIRVAIKKISPEHQTYCQRTLEIKILLRFRHENIIGINDILRAR 124
QY 111 TLEAMRDVYIQDLMETDLYKLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPS 170
DB 125 RIYEMRDVYIQDLMETDLYKLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPS 184
QY 171 NLLINTTCDLKI CDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSIDIS 230
DB 185 NLLINTTCDLKI CDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTSIDIS 244
QY 231 VGCILAEMLSNRPIPGKHVLDQNLHILGILSPQSDQEDLNCIINNKARNYLOSLSKTKV 290
DB 245 VGCILAEMLSNRPIPGKHVLDQNLHILGILSPQSDQEDLNCIINNKARNYLOSLSKTKV 304
QY 291 AWAKLFPKSDSKALDLDLDMRLTFNPNKRIT-----VAEPPFTFAM 330
DB 305 PNWKLFPKADNKALDLDLDMRLTFNPNKRITVEALAHPLYEQYDPTDEPVAEPPFTFAM 364
QY 331 ELDDLKERLKELIPOETARFQ 352
DB 365 ELDDLKERLKELIPOETARFQ 386

RESULT 6
JQ1400


```
Db 1 MAAAAASSPFGGP-----EMVRQAFDVGPRYTNLSYIGEGAYGMVCSAH 46
QY 61 DHVKTTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 47 CNINKVRVAIKKISPFHQYTCORTLREIKILLRFRKHENIIGINDIIRAPTIQMKDVYI 106
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 107 VQDLMETDLYKLLKTOHLSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 166
QY 181 KICDFGLARIADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
Db 167 KICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 226
QY 241 NRPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSTKVAVAKLFPKSD 300
Db 227 NRPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLLSLPHKNKVPWNLFPNAD 286
QY 301 SKALDILDRMLTFNPKRI-----TVAEPPFTFAMELDDLPERLK 340
Db 287 PKALDILDRMLTFNPKRIEVEAALAHPLYEQYDPSDFVABAPLKFEMELDDLPERLK 346
QY 341 KELIFOETARFQPG 354
Db 347 KELIFETARFQPG 360

RESULT 11
JW0053
extracellular signal-regulated kinase (EC 2.7.-.-) 2 - common carp
N;Alternate names: CERK2
C;Species: Cyprinus carpio (common carp)
C;Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C;Accession: JW0053
R;Hashimoto, H.; Yokoyama, Y.; Matsuo, Y.; Toyohara, H.; Kohno, M.; Sakaguchi, M.
J. Biochem. 123, 1031-1035, 1998
A;Title: Existence of two isoforms of extracellular signal-regulated kinase in fish.
A;Reference number: JW0052; MUID:98269030; PMID:9603989
A;Molecule type: mRNA
A;Residues: 1-369 <HAS>
A;Cross-references: UNIPROT:O7LZH5; UNIPARC:UPI0000175597
C;Comment: This enzyme is the key enzyme in the intracellular signal transduction pathway
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase
F;32-322/Domain: protein kinase homology <KIN>

Query Match 80.3%; Score 1508; DB 2; Length 369;
Best Local Similarity 77.7%; Pred. No. 4.7e-62;
Matches 290; Conservative 21; Mismatches 32; Indels 30; Gaps 3;

QY 2 AAAAAQGGGGEPRTTEGVGPGVGEVEMVKGQDFDVGPRYTOLOYIGEGAYGMVSSAYD 61
Db 4 AAVSAPAGGGPNP-----GSG-----AEMVRQAFDVGPRYTNLSYIGEGAYGMVCSAYK 53
QY 62 HVKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 121
Db 54 RDNKVRVAIKKISPFHQYTCORTLREIKILLRFRKHENIIGINDIIRTTIDQMDVYI 113
QY 122 QDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLK 181
Db 114 QDLMETDLYKLLKTOHLSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLK 173
QY 182 ICDGLARIADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 241
Db 174 ICDGLARVADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 233
QY 242 RPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSTKVAVAKLFPKSDS 301
Db 234 RPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLLSLPHKNKVPWNLFPNADP 293
QY 302 KALDILDRMLTFNPKRI-----TVAEPPFTFAMELDDLPERLK 341
```

```
Db 294 KALDILDRMLTFNPKRIEVEAALAHPLYEQYDPTDEPVABAPKFKFMELDDLPERLK 353
QY 342 ELIFOETARFQPG 354
Db 354 ELIFETARFQPG 366

RESULT 12
A46036
extracellular signal-regulated kinase (EC 2.7.1.1-) DmERK-A - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46036
R;Biggs III, W.H.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 89, 6295-6299, 1992
A;Title: Primary structure, expression, and signal-dependent tyrosine phosphorylation of
A;Reference number: A46036; MUID:92335284; PMID:1378625
A;Accession: A46036
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-376 <BIG>
A;Cross-references: UNIPROT:P40417; UNIPARC:UPI0000000FB3; GB:M95124; NID:9485754; PIDN:I
A;Note: sequence extracted from NCBI backbone (NCBIN:108684, NCBI:P:108685)
C;Genetics:
A;Gene: FlyBase:rl
A;Cross-references: FlyBase:FBgn0003256
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F;36-326/Domain: protein kinase homology <KIN>
F;44-52/Region: protein kinase ATP-binding motif

Query Match 72.6%; Score 1364.5; DB 2; Length 376;
Best Local Similarity 74.9%; Pred. No. 1.5e-55;
Matches 262; Conservative 28; Mismatches 39; Indels 21; Gaps 2;

QY 24 VP-GEVEMVKGQDFDVGPRYTOLOYIGEGAYGMVSSAYDHVKTTRVAIKKISPFHQYTC 82
Db 19 VPQSNAEVIRGQIFEVGPRYIKLAVIGEGAYGMVVSADDTLTNQVRVAIKKISPFHQYTC 78
QY 83 QRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIQQDLMETDLYKLLKSQQLSNDH 142
Db 79 QRTLREITLTFRKHENIIDIRLDVDSIDQMDRVYIVQCLMETDLYKLLKTQRLSNDH 138
QY 143 ICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPDHDTGFLT 202
Db 139 ICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPDHDTGFLT 198
QY 203 EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQNLHILGILG 262
Db 199 EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQNLHILGVLG 258
QY 263 SPSQEDLNCIINMKARNYLOSLPSTKVAVAKLFPKSDSKALDILDRMLTFNPKRI--- 319
Db 259 SPSRDLECIINMKARNYLESPLFPKPNVPWAKLFPFNADALADLLGKMLTFNPKRIPVE 318
QY 320 -----TVAEPPFTFAMELDDLPERLKELFOETARFQ 352
Db 319 EALAHPLYEQYDGPDEPVAEVFPFRINMENDDISRDALKSLIFEETLKF 368

RESULT 13
A36978
MAP kinase mpk-1 (EC 2.7.1.1-) - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: A36978; T22118
R;Lackner, M.R.; Kornfeld, K.; Miller, L.M.; Horvitz, H.R.; Kim, S.K.
Genes Dev. 8, 160-173, 1994
A;Title: A MAP kinase homolog, mpk-1, is involved in ras-mediated induction of vulval cel
A;Reference number: A36978; MUID:94131270; PMID:8299936
A;Accession: A36978
A;Status: preliminary
```

A:Molecule type: mRNA
A:Residues: 1-376 <LAC>
A:Cross-references: UNIPROT:P39745; UNIPARC:UPI000002B465; GB:U27124; GB:S68854; NID:g86
R:Jassal, B.; Smith, A
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19519
A:Accession: T22118
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-376 <WIL>
A:Cross-references: UNIPARC:UPI000002B465; EMBL:Z46937; PIDN:CAA87057.1; GSPDB:GN00021;
A:Experimental source: clone F43C1
C:Genetics:
A:Gene: mpk-1
A:Map position: 3
A:Introns: 23/3; 43/2; 85/3; 104/2; 304/2
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:26-285/Domain: protein kinase homology <KIN>
F:34-42/Region: protein kinase ATP-binding motif

Query Match 70.1%; Score 1317; DB 2; Length 376;
Best Local Similarity 72.4%; Pred. No. 2.2e-53;
Matches 246; Conservative 36; Mismatches 36; Indels 20; Gaps 1;

QY 28 VEMVKGQPFVGVPRYTQIQYIGEGAYGMVSSAYDHRVAKTRVAIKKISPFPHQTYCQRTL 87
DB 14 VEEVHGQLFVAPRYVNLISYIGEGAYGMVASALDTITRDRVAIKKISPFPHQTYCQRTL 73
QY 88 EIQILLRFPHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSKQSLNDHICYFL 147
DB 74 EIKILNRFKHENIINQIIRSETVDSIKDIYVQCLMETDLYKLKSKQSLNDHICYFL 133
QY 148 YQILRGKLYHSANVLRDLKPSNLLINTTCDLKICDFGLARVADPEHDHTGFLTEYVAT 207
DB 134 YQILRGKLYHSANVLRDLKPSNLLINTTCDLKICDFGLARVADPEHDHTGFLTEYVAT 193
QY 208 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDQNLHILGILGSPSQE 267
DB 194 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDQNLHILGILGSPSQE 253
QY 268 DLNCIINMKARNYLOSLPSKTKVAVAKLPKSDSKALDLDRLMLTFNPNKRI----- 319
DB 254 DLQCIINDKARSYLSLPHKPKQWARLYPGADPRALDLDLKMFLTNPNNRIDIEQALAH 313
QY 320 -----TVAREPFTFAMELDDLPKELKELIFQ 347
DB 314 PYLEQYDPGDEPVCBEPFTEMEFDDLPKELKELIWE 353

RESULT 14
A36977
MAP kinase sur-1 (EC 2.7.1.1) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A36977
R:Wu, Y.; Han, W.
Genes Dev. 8, 147-159, 1994
A:Title: Suppression of activated Let-60 ras protein defines a role of Caenorhabditis el
A:Reference number: A36977; MUID:94131269; PMID:8299935
A:Accession: A36977
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-444 <WUA>
A:Cross-references: UNIPROT:P39745; UNIPARC:UPI0000164181; GB:U03879; NID:g433356; PIDN:
C:Genetics:
A:Gene: sur-1
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:94-353/Domain: protein kinase homology <KIN>
F:102-110/Region: protein kinase ATP-binding motif

Query Match 70.1%; Score 1317; DB 2; Length 444;

Best Local Similarity 72.4%; Pred. No. 2.5e-53;
Matches 246; Conservative 38; Mismatches 36; Indels 20; Gaps 1;

QY 28 VEMVKGQPFVGVPRYTQIQYIGEGAYGMVSSAYDHRVAKTRVAIKKISPFPHQTYCQRTL 87
DB 82 VEEVHGQLFVAPRYVNLISYIGEGAYGMVASALDTITRDRVAIKKISPFPHQTYCQRTL 141
QY 88 EIQILLRFPHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSKQSLNDHICYFL 147
DB 142 EIKILNRFKHENIINQIIRSETVDSIKDIYVQCLMETDLYKLKSKQSLNDHICYFL 201
QY 148 YQILRGKLYHSANVLRDLKPSNLLINTTCDLKICDFGLARVADPEHDHTGFLTEYVAT 207
DB 202 YQILRGKLYHSANVLRDLKPSNLLINTTCDLKICDFGLARVADPEHDHTGFLTEYVAT 261
QY 208 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDQNLHILGILGSPSQE 267
DB 262 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDQNLHILGILGSPSQE 321
QY 268 DLNCIINMKARNYLOSLPSKTKVAVAKLPKSDSKALDLDRLMLTFNPNKRI----- 319
DB 322 DLQCIINDKARSYLSLPHKPKQWARLYPGADPRALDLDLKMFLTNPNNRIDIEQALAH 381
QY 320 -----TVAREPFTFAMELDDLPKELKELIFQ 347
DB 382 PYLEQYDPGDEPVCBEPFTEMEFDDLPKELKELIWE 421

RESULT 15
T51944
pathogenicity MAP kinase 1 [imported] - Pyricularia grisea
C:Species: Pyricularia grisea
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T51944
R:Xu, J.R.; Hamer, J.E.
Genes Dev. 10, 2696-2706, 1996
A:Title: MAP kinase and CAMP signaling regulate infection structure formation and pathoge
A:Reference number: Z25880; MUID:97102671; PMID:8946911
A:Accession: T51944
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-356 <XU>
A:Cross-references: UNIPROT:Q92246; UNIPARC:UPI000006C824; EMBL:U70134; PIDN:AAC49521.2
A:Experimental source: strain Guy11
C:Genetics:
A:Gene: PMK1
A:Introns: 39/2; 232/1
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 49.6%; Score 932; DB 2; Length 356;
Best Local Similarity 53.4%; Pred. No. 5.8e-36;
Matches 181; Conservative 55; Mismatches 79; Indels 24; Gaps 2;

QY 36 FDVGPRTYQIQYIGEGAYGMVSSAYDHRVAKTRVAIKKISPFPHQTYCQTLREIQILLRF 95
DB 18 FNVSEQYDIQDVVGEAYGVVCSAIHKPSQKQVAIKKITPFDHSMFCLETLREMKLRYF 77
QY 96 RHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSKQSLNDHICYFLYQILGLK 155
DB 78 NHENISILDIQKPRSVETNEVYLIQELMETDHRVIRVTDLSDDHCQFYIQTALAK 137
QY 156 YIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEI 215
DB 138 AMHSANVLRDLKPSNLLNANCDLVKDFGLARSAAQENNSGFMTYVATRWYRAPEI 197
QY 216 MLNSGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDQNLHILGILGSPSQEDLNCIINM 275
DB 198 MLTFKTYTKAIDVWSVGCILAEMLSNGKPLFPKGKDYHQLTLILDVLGTPTWEDYIGKSR 257
QY 276 KARNYLOSLPSKTKVAVAKLPKSDSKALDLDRLMLTFNPNKRIVAE----- 323
DB 258 RAREYINSLFPKKVPRTLPFKTSLDLDLLEKLAFAFPVKRITVEALKHPYLEPYHD 317

QY 324 ---EP-----FTFAMELDDLPKERLKEIFQETAR 350
Db 318 PDDEPTAPIPEEFDKHKONLSKEQLXQFIYQEIIMR 356

Search completed: February 6, 2006, 15:35:34
Job time : 27.8991 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:20:01 ; Search time 171.223 Seconds
(without alignments)
1479.265 Million cell updates/sec

Title: US-10-623-108-4
Perfect score: 1879
Sequence: 1 MAAAAAQGGGGGPRRTGV.....LKELIFQTARFGVLEAP 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1859	98.9	379	1 MK03 HUMAN	P27361 h mitogen-a
2	1787	95.1	380	1 MK03 MOUSE	Q63844 h mitogen-a
3	1784	94.9	380	1 MK03 RAT	P21708 r mitogen-a
4	1784	94.9	380	2 OAPI78 RAT	Q4piV8 rattus norv
5	1767	94.0	369	2 Q7M0H9 CRIGR	Q7m0h9 cricetus
6	1683	89.6	357	2 Q8NHX1 HUMAN	Q8nhx1 homo sapien
7	1593	84.8	335	2 Q8NHX0 HUMAN	Q8nhx0 homo sapien
8	1561	83.1	323	2 Q7Z3H5 HUMAN	Q7z3h5 homo sapien
9	1551.5	82.6	392	2 Q7ZVK8 BRARE	Q7zvk8 brachydanio
10	1550.5	82.5	391	2 Q9DGR6 BRARE	Q9dgr6 brachydanio
11	1550.5	82.5	392	2 Q7LZH4 CYPCA	Q7lzh4 cyprinus ca
12	1538.5	81.9	360	1 MK01 BOVIN	P46196 bos taurus
13	1538.5	81.9	360	1 MK01 HUMAN	P28482 homo sapien
14	1537	81.8	368	2 Q8UMG6 CHICK	Q8umg6 gallus gall
15	1528.5	81.3	358	1 MK01 MOUSE	P63085 mus musculus
16	1528.5	81.3	358	1 MK01 RAT	P63085 rattus norv
17	1527	81.3	358	2 Q8NVR1 PONPY	Q8nvr1 pongo pygma
18	1522	81.0	361	1 MK01 XENLA	P26696 xenopus lae
19	1520	80.9	361	2 Q5D0K1 XENLA	Q5d0k1 xenopus lae
20	1511	80.4	369	2 Q6P0Z3 BRARE	Q6p0z3 brachydanio
21	1509	80.3	361	2 Q6DFK6 XENLA	Q6dfk6 xenopus lae
22	1508	80.3	369	2 Q7LZH5 CYPCA	Q7lzh5 cyprinus ca
23	1506	80.1	369	2 Q5I811 CYPCA	Q5i811 cyprinus ca
24	1504	80.0	369	2 Q7ZW72 BRARE	Q7zw72 brachydanio
25	1501	79.9	369	2 Q9DGR5 BRARE	Q9dgr5 brachydanio
26	1495	79.6	369	2 Q4VH14 BRARE	Q4vhl14 brachydanio
27	1488	79.2	358	2 Q4S9M7 TETNG	Q4s9m7 tetraodon n
28	1488	79.2	366	2 Q4RV04 TETNG	Q4rv04 tetraodon n
29	1445	76.9	351	2 Q16945 AFLCA	Q16945 alysia cal
30	1418	75.5	369	2 Q8IS92 STRPU	Q8is92 strongyloce
31	1399	74.5	350	2 Q4H3L5 CIOIN	Q4h3l5 ciona intes

32	1384.5	73.7	395	2 Q8AV10 PAROL	Q8avi0 parolichthy
33	1372.5	73.0	372	2 Q86N95 MARGL	Q86n95 marthasteri
34	1371.5	73.0	362	2 Q8MXZ5 HALRO	Q8mxz5 halocynthia
35	1364.5	72.6	376	1 ERKA DROME	P40417 drosophila
36	1364.5	72.6	376	2 Q7PL59 DROME	Q7pl59 drosophila
37	1355	72.1	394	2 Q5K4T9 9BILA	Q5k4t9 meloidogyne
38	1323	70.4	374	2 Q6ORP3 CABER	Q6orp3 caenorhabdi
39	1317	70.1	444	1 SURI CAEEL	P39745 caenorhabdi
40	1317	70.1	444	1 SURI CAEEL	P39745 caenorhabdi
41	1242.5	66.1	291	2 Q35558 MOUSE	Q35558 mus musculus
42	1163.5	61.9	351	2 Q6PNM5 SCHJA	Q6pnm5 schistosoma
43	1095	58.3	230	2 Q7PVR1 ANOGA	Q7pvr1 anopheles g
44	944	50.2	352	2 Q876L7 LEPMC	Q876l7 leptosphaer
45	944	50.2	352	2 Q96X31 9PLEO	Q96x31 pyrenophora
46	944	50.2	352	2 Q6R3Q3 ALTER	Q6r3q3 alternaria

ALIGNMENTS

RESULT 1
MK03_HUMAN ID MK03_HUMAN STANDARD; PRT; 379 AA.
AC P27361;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
DE associated protein 2 kinase).
GN Name=MAPK3; Synonyms=ERK1, PRK3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hepatoma;
RX MEDLINE=93330262; PubMed=7687743;
RA Charest D.L., Jirik F., Harder K., Pelech S.L., Mordret G.;
RT "Molecular cloning, expression, and characterization of the human
RL mitogen-activated protein kinase p44erk1.";
RL Mol. Cell. Biol. 13:4679-4690(1993).
RN [2]
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RC NUCLEOTIDE SEQUENCE OF 14-379.
RX MEDLINE=92171961; PubMed=1540184;
RA Owaki H., Makar R., Boulton T.G., Cobb M.H., Geppert T.D.;
RT "Extracellular signal-regulated kinases in T cells: characterization

RT of human ERK1 and ERK2 cDNAs.";
 RL Biochem. Biophys. Res. Commun. 182:1416-1422 (1992).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 25-379.
 RX MEDLINE=92316223; PubMed=1319925; DOI=10.1016/0014-5793(92)80612-K;
 RA Gonzalez F.A., Raden D.L., Rigby M.R., Davis R.J.;
 RT "Heterogeneous expression of four MAP kinase isoforms in human
 tissues.";
 RL FEBS Lett. 304:170-178 (1992).
 RN [5]
 RP INTERACTION WITH HIV-1 NEF.
 RX PubMed=8794306;
 RA Greenway A.L., Azad A., Mills J., McPhee D.A.;
 RT "Human immunodeficiency virus type 1 Nef binds directly to LCK and
 RT Mitogen-activated protein kinase, inhibiting kinase activity.";
 RL J. Virol. 70:6701-6708 (1996).
 CC -I- FUNCTION: Involved in both the initiation and regulation of
 CC meiosis, mitosis, and postmitotic functions in differentiated
 CC cells by phosphorylating a number of transcription factors such as
 CC ELK-1. Phosphorylates E1F4EBP1; required for initiation of
 CC translation. Phosphorylates microtubule-associated protein 2
 CC (MAP2) (By similarity).
 CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -I- COFACTOR: Magnesium (By similarity).
 CC -I- ENZYME REGULATION: Activated and tyrosine phosphorylated in
 CC response to insulin and NGF.
 CC -I- SUBUNIT: Binds to HIV-1 Nef. This interaction inhibits its kinase
 CC activity.
 CC -I- INTERACTION:
 CC P53355:DAPK1; NbExp=3; IntAct=EBI-73995; EBI-358616;
 CC O75676:RPS6KA4; NbExp=1; IntAct=EBI-73995; EBI-73933;
 CC -I- PTM: Autophosphorylated on threonine and tyrosine residues (By
 CC similarity).
 CC -I- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
 CC kinase subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; X60188; CAA42744.1; -; mRNA.
 DR EMBL; BC013992; ABH13992.1; -; mRNA.
 DR EMBL; M84490; AAA36142.1; -; mRNA.
 DR EMBL; Z11696; CAA77754.1; -; mRNA.
 DR FIR; A48082; A48082.
 DR HSSP; P28482; 1PME.
 DR PIR; P27361; 29-374.
 DR SWR; P27361; 29-374.
 DR IntAct; P27361; -;
 DR Ensembl; ENSG00000102882; Homo sapiens.
 DR HGNC; HGNC:6877; MAPK3.
 DR H-InvDB; HIX0012930; -;
 DR Reactome; P27361; -;
 DR MIM; 601795; -;
 DR GO; GO:0005524; F-ATP binding; NAS.
 DR GO; GO:0004707; F-MAP kinase activity; NAS.
 DR GO; GO:000468; P:protein amino acid phosphorylation; IDA.
 DR GO; GO:000074; P:regulation of cell cycle; NAS.
 DR InterPro; IPR008349; Erk 1 2 MAPK.
 DR InterPro; IPR008351; JNK/MAPK.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; Pkinase; I.
 DR PRINTS; PR01770; ERK1ERK2MAPK.
 DR PRINTS; PR01772; JNK/MAPKINASE.
 DR ProDom; PD000001; Prot_KINASE.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 42 330 Protein kinase.
 FT NP_BIND 48 56 ATP (By similarity).
 FT ACT_SITE 166 166 Proton acceptor (By similarity).
 FT BINDING 71 71 ATP (By similarity).
 FT MOD_RES 202 202 Phosphothreonine (activates the kinase).
 FT MOD_RES 204 204 Phosphotyrosine (activates the kinase).
 FT CONFLICT 174 174 I -> S (in Ref. 1).
 SQ SEQUENCE 379 AA; 43136 MW; E6020CE413EC41F7 CRC64;
 Query Match 98.9%; Score 1859; DB 1; Length 379;
 Best Local Similarity 94.7%; Pred No. 4.7e-110;
 Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
 QY 1 MAAAAAQQGGGGEPRRTGEGVGVPEVEMVKGQPDFVGPRTYQIYIGEGAYGMVSSAY 60
 DB 1 MAAAAAQQGGGGEPRRTGEGVGVPEVEMVKGQPDFVGPRTYQIYIGEGAYGMVSSAY 60
 QY 61 DHVTRTRVAIKKISPFHQTYCQRTIREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
 DB 61 DHVTRTRVAIKKISPFHQTYCQRTIREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
 QY 121 VDLMETDLYKLLKSQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 DB 121 VDLMETDLYKLLKSQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDIWSVCILAEMLLS 240
 DB 181 KICDFGLARIADPEHDHTGFLTYVATRWYRAPEIMLSKGYTKSIDIWSVCILAEMLLS 240
 QY 241 NRPIPGKHYLDQNLNHIILGILGSPQEDLNCKINKARNYLOSLSKTKVAKLPPKSD 300
 DB 241 NRPIPGKHYLDQNLNHIILGILGSPQEDLNCKINKARNYLOSLSKTKVAKLPPKSD 300
 QY 301 SKALDLDRLMFLPNFKRIT-----VAEEPTFAMELDDLPKERL 340
 DB 301 SKALDLDRLMFLPNFKRITVEALAHPLYEQYDPTDEPVAEEPTFAMELDDLPKERL 360
 QY 341 KELIFQETARFQGVLEAP 359
 DB 361 KELIFQETARFQGVLEAP 379
 RESULT 2
 MK03_MOUSE
 ID MK03_MOUSE STANDARD; PRT; 380 AA.
 AC Q63844; Q61531; Q8K0X5; Q91YW5;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
 DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
 DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (BRT2) (p44-MAPK) (Microtubule-
 DE associated protein 2 kinase) (MNK1).
 GN Name-Mpk3; Synonyms-Erk1, Prkm3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 OX [1]_TaxID=10090;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP STRAIN=FVB/N; TISSUE=Kidney, and Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,


```

OY 342 ELIFQETARFQGVLEAP 359
Db 363 ELIFQETARFQGVLEAP 380

RESULT 3
MK03 RAT STANDARD; PRT; 380 AA.
AC P21708; Q62686; Q9J1J3;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 1) (ERK1-1) (Insulin-stimulated MAP2 kinase)
DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
DE associated protein 2 kinase) (MNK1).
GN Names=Mapk3; Synonyms=Erk1, Prkm3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=93013050; PubMed=1327976; DOI=10.1016/0378-1119(92)90109-3;
RA Marquardt B., Stabel S.;
RT "Sequence of a rat cDNA encoding the ERK1-MAP kinase.";
RL Gene 120:297-299(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 7-380.
RA Maisonnier P.C., le Beau M.M., Espinosa R. III, Ip N.Y.
RA Belluscio L., la Monte S.M., Squinco S., Furth M.E., Yancopoulos G.D.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE OF 14-380, AND PARTIAL PROTEIN SEQUENCE.
RX MEDLINE=90312137; PubMed=2164259;
RA Boulton T.G., Yancopoulos G.D., Gregory J.S., Slaughter C., Moomaw C.,
RA Hsu J., Cobb M.H.;
RT "An insulin-stimulated protein kinase similar to yeast kinases
RT involved in cell cycle control.";
RL Science 249:64-67(1990).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 14-380.
RC TISSUE=Brain cortex;
RX MEDLINE=91369479; PubMed=1716439;
RA de Miguel C., Kligman D., Patel J., Detera-Wadleigh S.D.;
RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
RT from mouse and rat brain.";
RL DNA Cell Biol. 10:505-514(1991).
RN [5]
RP NUCLEOTIDE SEQUENCE (ISOFORM B).
RX MEDLINE=20283605; PubMed=10748187; DOI=10.1074/jbc.M910060199;
RA Yung Y., Yao Z., Hancock T., Seger R.;
RT "ERK1b, a 46-kDa ERK isoform that is differentially regulated by
RT MEK";
RL J. Biol. Chem. 275:15799-15808(2000).
RN [6]
RP PROTEIN SEQUENCE OF 43-64 AND 167-185, AND CHARACTERIZATION.
RX MEDLINE=91105092; PubMed=1845291;
RA Boulton T.G., Gregory J.S., Cobb M.H.;
RT "Purification and properties of extracellular signal-regulated kinase
RT 1, an insulin-stimulated microtubule-associated protein 2 kinase.";
RL Biochemistry 30:278-286(1991).
RN [7]
RP AUTOPHOSPHORYLATION.
RX MEDLINE=91296777; PubMed=1712480;
RA Seger R., Ahn N.G., Boulton T.G., Yancopoulos G.D., Panayotatos N.,
RA Radziejewska E., Ericsson L., Bratlien R.L., Cobb M.H., Krebs E.G.;
RT "Microtubule-associated protein 2 kinases, ERK1 and ERK2, undergo
RT autophosphorylation on both tyrosine and threonine residues:
RT implications for their mechanism of activation.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:6142-6146(1991).

```

```

RN RP PHOSPHORYLATION OF E1F4BP1.
RX MEDLINE=95025978; PubMed=7939721;
RA Lin T.-A., Kong X., Haystead T.A.J., Pause A., Belsham G.J.,
RA Sonenberg N., Lawrence J.C. Jr.;
RT "PHAS-I as a link between mitogen-activated protein kinase and
RT translation initiation.";
RL Science 266:653-656(1994).
CC -!- FUNCTION: Involved in both the initiation and regulation of
CC meiosis, mitosis, and postmitotic functions in differentiated
CC cells by phosphorylating a number of transcription factors such as
CC ELK-1. Phosphorylates E1F4BP1; required for initiation of
CC translation. Phosphorylates microtubule-associated protein 2
CC (MAP2) (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated and tyrosine phosphorylated in
CC response to insulin and NGF.
CC -!- SUBCELLULAR LOCATION: Nuclear (isoform B).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=P21708-1; Sequence=Displayed;
CC Name=B;
CC IsoId=P21708-2; Sequence=VSP_004830;
CC -!- TISSUE SPECIFICITY: Highest levels within the nervous system,
CC expressed in different tissues, mostly in intestine, placenta and
CC lung.
CC -!- DEVELOPMENTAL STAGE: Increased expression during development.
CC -!- PTM: Autophosphorylated on threonine and tyrosine residues.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
EMBL; X65198; CAA46318.1; -; mRNA.
EMBL; M61177; AAA63486.1; -; mRNA.
EMBL; M38194; AAA41123.1; -; mRNA.
EMBL; U12008; AAA20009.1; -; mRNA.
EMBL; AF155236; AAF71666.1; -; mRNA.
PIR; JC1451; JC1451.
HSP; P28482; LPWE.
SMR; P21708; 30-375.
Ensembl; ENSRNOG00000019601; Rattus norvegicus.
RGD; 3046; Mapk3.
GO; GO:0005829; C:cytosol; IDA.
GO; GO:0005626; C:insoluble fraction; IDA.
GO; GO:0004707; F:MAP kinase activity; TAS.
GO; GO:0000189; P:nuclear translocation of MAPK; IDA.
GO; GO:0045941; P:positive regulation of transcription; IE.
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GO; GO:0007243; P:protein kinase cascade; TAS.
InterPro; IPR008349; Erk_1-2_MAPK.
InterPro; IPR008351; JNK_MAPK.
InterPro; IPR003527; MAP_kin.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR008271; Ser_thr_kin_AS.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; Pkinase; 1.
PRINTS; PR01770; ERK1ERK2MAPK.
PRINTS; PR01772; JNKMAPKINASE.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TK; 1.
PROSITE; PS01351; MAPK; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS50108; PROTEIN_KINASE_ST; 1.
KW Alternative splicing; ATP-binding; Cell cycle;
KW Direct protein sequencing; Kinase; Nuclear protein;

```



```
DR Pfam: PF00069; Pkinase: 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase: 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_SF; 1.
DR ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 369 369
FT NON_TER 1 1
SQ SEQUENCE 369 AA; 42126 MW; 0BB8FCAAC690B4FF CRC64;

Query Match 94.0%; Score 1767; DB 2; Length 369;
Best Local Similarity 92.4%; Pred. No. 3.2e-104;
Matches 341; Conservative 1; Mismatches 7; Indels 20; Gaps 1;

QY 11 GGEPRTEGVGPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAYDHRVTRVAI 70
Db 1 GGEPRGAAGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAYDHRVTRVAI 60
QY 71 KKISPFHQTYCQRTLRQIILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLY 130
Db 61 KKISPFHQTYCQRTLRQIILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLY 120
QY 131 KLLKSQSLNDHICYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDLKI CDFGLARI 190
Db 121 KLLKSQSLNDHICYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDLKI CDFGLARI 180
QY 191 ADEPHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSNRPFPKGY 250
Db 181 ADEPHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSNRPFPKGY 240
QY 251 LDQLNHILGLSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLLDRM 310
Db 241 LDQLNHILGLSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLLDRM 300
QY 311 LTFNPNKRIT-----VAREPTFAMELDDLPKRLKELI FOETAR 350
Db 301 LTFNPNKRITVEDALAHPLYEQYDPTDSEPVAREPTFDMELDDLPKRLKELI FOETAR 360
QY 351 FQPGVLEAP 359
Db 361 FQGAPEAP 369

RESULT 6
Q8NHX1 HUMAN
ID Q8NHX1_HUMAN PRELIMINARY; PRT; 357 AA.
AC Q8NHX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-related kinase 1b.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Aebersold D.N., Yung Y., Seger R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033607; AAK52329.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q8NHX1; 29-344.
DR Ensembl; ENSG00000102882; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.

DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase: 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
DR ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 357 AA; 40062 MW; 58C92773983ADA79 CRC64;

Query Match 89.6%; Score 1683; DB 2; Length 357;
Best Local Similarity 99.4%; Pred. No. 6.5e-99;
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAAAAQQGGGEGPRTEGVGPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGEGPRTEGVGPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
QY 61 DHVTRTRVAIKKISPFHQTYCQRTLRQIILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVTRTRVAIKKISPFHQTYCQRTLRQIILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VDLMETDLYKLLKSQSLNDHICYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDL 180
Db 121 VDLMETDLYKLLKSQSLNDHICYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQLNHILGLSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLFPKSD 300
Db 241 NRPIFGKHYLDQLNHILGLSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLFPKSD 300
QY 301 SKALDLLDRMLTFNPNKRITVAE 323
Db 301 SKALDLLDRMLTFNPNKRITVEE 323

RESULT 7
Q8NHX0 HUMAN
ID Q8NHX0_HUMAN PRELIMINARY; PRT; 335 AA.
AC Q8NHX0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-related kinase 1c.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Aebersold D.M., Yung Y., Seger R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033608; AAK52330.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q8NHX0; 29-330.
DR Ensembl; ENSG00000102882; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
```


RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB; TISSUE=Whole body;
RG NIH MGC Project;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Director MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tuebingen;
RA Krens S.F.G., Snaar-Jagalska B.E., Spaink H.P.;
RT "Zebrafish ERK1 and ERK2.";
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045505; AAH45505.1; -; mRNA.
DR EMBL; BC097073; AAH97073.1; -; mRNA.
DR EMBL; BC066401; AAH66401.1; -; mRNA.
DR EMBL; AY922319; AAY57804.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SNR; Q7ZVK8; 41-386.
DR ZFIN; ZDB-GENE-040121-1; mapk3.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK_MAPK.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKG; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 392 AA; 44099 MW; C184681E448E5B8E CRC64;

Query Match 82.6%; Score 1551.5; DB 2; Length 392;
Best Local Similarity 79.9%; Pred. No. 1.6e-90;
Matches 303; Conservative 15; Mismatches 32; Indels 29; Gaps 3;

QY 2 AAAAAAG-----GGCGEPRRTGCVGPGEVVMVKGPFDVGPRTYQYICEGAY 53
DB 9 AAAGAGAGNSAAGCGGAVAGCGPSGAAGSKPG-LESVKGQNFVGPRTYDLYIGEGAY 67
QY 54 GVVSSAYDHRVTRVAIKKISPFHQTYCQRTLEIRIQILLRPHRHENVIGIRDLRASTLE 113
DB 68 GWCVSFAFNVNKKIRVAIKKISPFHQTYCQRTLEIRIKILLRPHRHENIIGINDILRAHID 127
QY 114 ANRDYIVQDLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLL 173
DB 128 YNRDYYIVQDLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLL 187
QY 174 INTTCDLKICDPLGARIADPEHDHGTGFLTEYVATRWYRAPEIMLNSKGYTSIDISWVGC 233
DB 188 INTTCDLKICDPLGARIADPEHDHGTGFLTEYVATRWYRAPEIMLNSKGYTSIDISWVGC 247
QY 234 ILAEMLSNRPIPPGKHLYDLQNLHILGILGSPSQEDLNCIINNKARNYLQSLPSKTKVANA 293
DB 248 ILAEMLSNRPIPPGKHLYDLQNLHILGILGSPSQEDLNCIINNKARNYLQSLPSKTKIPWN 307
QY 294 KLPFKSDSKALDLDRLMTFNPKNKIT-----VAEEPTFAMELD 333
DB 308 KLPFKADNKALDLDRLMTFNPKNKINVEQALAHPLYEQYDPSDEPVAEEPTFAMELD 367
QY 334 DLPKERLKLIFQETARFQ 352
DB 368 DLPKERLKLIFQETARFQ 386
RESULT 10
ID Q9DGR6_BRARE PRELIMINARY; PRT; 391 AA.
AC Q9DGR6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ERK1.
GN Name=mapk3; Synonyms=zERK1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hirano T., Fujii R., Hibi M.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030902; BAB11812.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q9DGR6; 40-385.
DR ZFIN; ZDB-GENE-040121-1; mapk3.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK_MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKG; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 391 AA; 44126 MW; 6638F0807B5F9889 CRC64;
Query Match 82.5%; Score 1550.5; DB 2; Length 391;

Best Local Similarity 81.1%; Pred. No. 1.8e-90;
Matches 300; Conservative 17; Mismatches 32; Indels 21; Gaps 2;

QY 3 AAAAGGGGPRRTSGVGPVGEVEMVKGQPDVGPVYQYIGEGAGVSSAYDH 62
DB 17 SSAAGPGGAVAPGGGAGSGKPG-LESVKGQDFVGPRTYDQYIGEGAGVCSAFDN 75

QY 63 VRKTRVAIKKISPFHQYQORTLREIQILRRPHENVIGIRDILRASTLEAMRDVYIVQ 122
DB 76 VNKIRVAIKKISPFHQYQORTLREIKILLRPHENIIGINDILRARHIDYMRDVYIVQ 135

QY 123 DLMETDLYKLLKSQQLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI 182
DB 136 DLMETDLYKLLKTQQLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI 195

QY 183 CDGLARIADPEHDHGTFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR 242
DB 196 CDGLARIADPEHDHGTFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR 255

QY 243 PIFPGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKWAKLFPKSDSK 302
DB 256 PIFPGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKWAKLFPKADNK 315

QY 303 ALDLDRLMTFNPVKRIT-----VAEEPTFAMELDDLPKRLKE 342
DB 316 ALDLDRLMTFNPVKRINVRQALAHPLYEQYDPSDPAEAEPTTFNMELDDLPKRLKE 375

QY 343 LIPOETARFQ 352
DB 376 LIPEETARFQ 385

RESULT 11
Q7LZH4_CYPCA PRELIMINARY; PRT; 392 AA.

ID Q7LZH4; AC Q7LZH4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Extracellular signal-regulated kinase (EC 2.7.-.-) 1 (Extracellular
DE signal regulated protein kinase 1).
GN Name=cERK1;
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=98269030; PubMed=9603989;
RX Hashimoto H., Yokoyama Y., Matsuo Y., Toyohara H., Kohno M.,
RA Sakaguchi M.;
RT "Existence of two isoforms of extracellular signal-regulated kinase in
RT fish.";
RL J. Biochem. 123:1031-1035(1998).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB006038; BAD23842.1; -, mRNA.
DR PIR; JW0052; JW0052.
DR HSSP; Q16539; 1021.
DR SMR; Q7LZH4; 39-386.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR003527; MAP Kin.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00069; Ekinase; I.
DR PRINTS; PR01770; ERKIERK2MAPK.
DR ProDom; PD000001; Prot kinase; 2.
DR PROSITE; PS01351; MAPK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 392 AA; 44133 MW; 02EE901325C347B CRC64;

Query Match 82.5%; Score 1550.5; DB 2; Length 392;
Best Local Similarity 78.5%; Pred. No. 1.8e-90;
Matches 300; Conservative 20; Mismatches 29; Indels 33; Gaps 3;

QY 3 AAAAGGGGPRRTSGVGPV-----PGEVEMVKGQPDVGPVYQYIGEGAGVSSAYDH 50
DB 6 SSAAGPGGAVAPGGGAGSGKPG-LESVKGQDFVGPRTYDQYIGEGAGVCSAFDN 64

QY 51 GAYGVSSAYDHVRKTRVAIKKISPFHQYQORTLREIQILRRPHENVIGIRDILRAS 110
DB 65 GAYGVCSAFDNVKNIRVAIKKISPFHQYQORTLREIKILLRPHENIIGINDILRAR 124

QY 111 TLEAMRDVYIVQDLMETDLYKLLKSQQLSNDHICVFLYQILRGLKYIHSANVLRDLKPS 170
DB 125 RIEYMRDVYIVQDLMETDLYKLLKTQQLSNDHICVFLYQILRGLKYIHSANVLRDLKPS 184

QY 171 NLLINTTCDLKI CDGLARIADPEHDHGTFLTEYVATRWYRAPEIMLSKGYTKSIDIWS 230
DB 185 NLLINTTCDLKI CDGLARIADPEHDHGTFLTEYVATRWYRAPEIMLSKGYTKSIDIWS 244

QY 231 VGCILAEMLSNRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKV 290
DB 245 VGCILAEMLSNRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKV 304

QY 291 AWAKLFPKSDSKALDLDRLMTFNPVKRIT-----VAEEPTFAMELDDLPKRLKE 330
DB 305 PNKLFPPKADNKALDLDRLMTFNPVKRITVBEALAHPLYEQYDPSDPAEAEPTTFNM 364

QY 331 ELDDLPKRLKE LIPOETARFQ 352
DB 365 ELDDLPKRLKE LIPEETARFQ 386

RESULT 12
MK01_BOVIN STANDARD; PRT; 360 AA.

ID MK01_BOVIN AC P46196;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
DE 2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERT1).
GN Name=MAPK1; Synonyms=ERK2, PRKM1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=adrenal medulla;
RA Ely C.M., Cox M.E., Her J., Parsons S.J.;
RT "Cloning and sequencing of ERK2 from a bovine adrenal medulla cDNA
RT library.";
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Phosphorylates microtubule-associated protein 2 (MAP2).
CC Myelin basic protein (MBP), and Elk-1; may promote entry in the
CC cell cycle.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
CC threonine in response to insulin and NGF.
CC -1- PTM: Autophosphorylated on threonine and tyrosine residues, which
CC correlates with a slow and low level of autoactivation (By
CC similarity).

```

CC CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC CC kinase subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: Z14089; CAA78467.1; -, mRNA.
CC PIR: S25011; S25011.
CC HSSP: P28482; 1PME.
CC SMR: P46196; 16-359.
CC InterPro: IPR008349; Erk_1_2_MAPK.
CC InterPro: IPR003527; MAP_kin.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR008271; Ser_thr_pkin_AS.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; Pkinase_1.
CC PRINTS: PR01770; ERK1ERK2MAPK.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS01351; MAPK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation;
KW Serine/threonine-protein kinase; Transferase.
FT DOMAIN 25 313
FT NP_BIND 31 39 ATP (By similarity).
FT COMPAS 2 9 Poly-Ala.
FT ACT_SITE 149 149 Proton acceptor (By similarity).
FT BINDING 54 54 ATP (By similarity).
FT MOD_RES 185 185 Phosphothreonine (activates the kinase)
FT MOD_RES 187 187 Phosphotyrosine (activates the kinase)
FT MOD_RES 187 187 Phosphoserine (By similarity)
FT MOD_RES 187 187 Phosphotyrosine (By similarity)
FT MOD_RES 187 187 Phosphoserine (By similarity)
SQ SEQUENCE 360 AA; 41376 MW; 58500B2A4E9549DE CRC64;

Query Match 81.9%; Score 1538.5; DB 1; Length 360;
Best Local Similarity 79.1%; Pred. No. 9.7e-90;
Matches 296; Conservative 18; Mismatches 23; Indels 37; Gaps 2;

QY 1 MAAAAAGGGGPRTEGVGPGVGEVEMVKQPFDPVGRYTOLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGAG-----PENVRGQVDFGVGRYTNLSYIGEGAYGMVCSAY 43
QY 61 DHVTRKTRVAIKLSPPEHOTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 44 DNVNKRVAIKLSPPEHOTYCORTLREIKILLRFRHENVIGINDIIRAPTIQMKDVYI 103
QY 121 VQDLMETDLYKLKLSQOLSNHDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTCDL 180
DB 104 VQDLMETDLYKLKLTQHLNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTCDL 163
QY 181 KICDFGLARIADPEHHTGTLTYVATRWYRAPEIMLSNGYTKSDIWSVGCILAEMLS 240
DB 164 KICDFGLARVADPDHDTGTLTYVATRWYRAPEIMLSNGYTKSDIWSVGCILAEMLS 223
QY 241 NRPIFGKHVLDQNLHLTGILGSPSQEDLNCIINMKARNYLOSLPKTKVAVAKLPKSD 300
DB 224 NRPIFGKHVLDQNLHLTGILGSPSQEDLNCIINMKARNYLSLPHKNKVPNNRLEFPNAD 283
QY 301 SKALDLDRLMTFNPKNRI-----TVASEPTFAMELDDLPKRL 340
DB 284 SKALDLDKMLTFNPKHRIEVAQALHPYLEQYDPSDEPVAAPKPFDMELDDLPKRL 343
QY 341 KELIFQETARFQPG 354
DB 344 KELIFETARFQPG 357

```

RESULT 13

```

MK01_HUMAN
ID MK01_HUMAN STANDARD; PRT; 360 AA.
AC P28482;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
DE 2) (MAP kinase 2) (MAPK 2) (P42-MAPK) (BRT1).
GN Name=MAPK1; Synonyms=ERK2, PRKMI;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OC NCBI_TaxID=9606;
RN [1]
RX MEDLINE=92171961; PubMed=1540184;
RA Owaki H., Makar R., Boulton T.G., Cobb M.H., Geppert T.D.;
RT "Extracellular signal-regulated kinases in T cells: characterization
RT of human ERK1 and ERK2 cDNAs."
RL Biochem. Biophys. Res. Commun. 182:1416-1422(1992).
RN [2]
RX NUCLEOTIDE SEQUENCE.
RP MEDLINE=92316223; PubMed=1319925; DOI=10.1016/0014-5793(92)80612-K;
RA Gonzalez F.A., Raden D.L., Ricby M.R., Davis R.J.;
RT "Heterogeneous expression of four MAP kinase isoforms in human
RT tissues."
RL FEBS Lett. 304:170-178(1992).
RN [3]
RX NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RX INTERACTION WITH HIV-1 NEF.
RX PubMed=8794306;
RA Greenway A.L., Azad A., Mills J., McPhee D.A.;
RT "Human immunodeficiency virus type 1 Nef binds directly to LCK and
RT mitogen-activated protein kinase, inhibiting kinase activity."
RL J. Virol. 70:6701-6708(1996).
CC -!- FUNCTION: Phosphorylates microtubule-associated protein 2 (MAP2).
CC Myelin basic protein (MBP), and Elk-1; may promote entry in the
CC cell cycle.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
CC threonine in response to insulin and NGF.
CC -!- SUBUNIT: Binds to HIV-1 Nef through its SH3 domain. This
CC interaction inhibits its tyrosine-kinase activity.
CC -!- PTM: Autophosphorylated on threonine and tyrosine residues, which
CC correlates with a slow and low level of autoactivation (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP

```

kinase subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; M84489; AAA58459.1; -; mRNA.
DR EMBL; Z11694; CAA7752.1; -; mRNA.
DR EMBL; Z11695; CAA7753.1; ALT_INIT; mRNA.
DR EMBL; BC017832; AAL17832.1; -; mRNA.
DR PIR; JQ1400; JQ1400.
DR PDB; 1PME; X-ray; @=1-360.
DR OGP; P28482; -.
DR Ensembl; ENSG00000100030; Homo sapiens.
DR HGNC; HGNC:6871; MAPK1.
DR H-InvDB; HIX0016281; -.
DR Reactome; P28482; -.
DR MIM; 176948; -.
DR GO; GO:0004707; F:MAP kinase activity; TAS.
DR GO; GO:0006935; P:chemotaxis; TAS.
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR GO; GO:0006950; P:response to stress; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR008349; Erk 1,2 MAPK.
DR InterPro; IPR003527; MAP_kin_
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Chtr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW 3D-structure; ATP-binding; Cell Cycle; Kinase; Nucleotide-binding;
KW Phosphorylation; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 25 313
FT NP_BIND 31 39
FT COMPBIAS 2 9
FT ACT_SITE 149 149
FT BINDING 54 54
FT MOD_RES 185 185
FT MOD_RES 187 187
FT CONFLICT 91 91
FT TURN 23 24
FT STRAND 25 30
FT STRAND 39 44
FT TURN 45 47
FT STRAND 50 56
FT TURN 59 60
FT TURN 62 77
FT TURN 81 82
FT STRAND 83 83
FT STRAND 88 90
FT TURN 95 97
FT STRAND 101 106
FT STRAND 110 111
FT HELIX 112 118
FT TURN 123 142
FT TURN 143 144
FT STRAND 145 146
FT TURN 152 154
FT STRAND 155 157
FT TURN 159 160
FT STRAND 163 165
FT TURN 169 170

FT STRAND 173 173
FT HELIX 176 178
FT STRAND 180 180
FT TURN 184 184
FT HELIX 191 193
FT HELIX 196 198
FT TURN 201 201
FT STRAND 202 202
FT TURN 207 207
FT HELIX 208 223
FT HELIX 233 244
FT HELIX 249 253
FT TURN 254 254
FT HELIX 258 265
FT TURN 266 266
FT HELIX 275 278
FT TURN 280 281
FT HELIX 284 293
FT TURN 298 300
FT HELIX 304 308
FT TURN 309 309
FT HELIX 311 313
FT TURN 314 316
FT HELIX 319 321
FT HELIX 340 350
FT HELIX 352 354
FT TURN 356 358
SQ SEQUENCE 360 AA; 41390 MW; E85D0B2A5D2D724E CRC64;
Query Match 81.9%; Score 1538.5; DB 1; Length 360;
Best Local Similarity 79.1%; Pred. No. 9.7e-90;
Matches 296; Conservative 18; Mismatches 23; Indels 37; Gaps 2;
QY 1 MAAAAAGGGGGPRRTGEGVGPVGEVEMVKQGPFDVGPRTYQLQYIGSGAYGMVSSAY 60
DB 1 MAAAAAGAG-----PEWVRQVDFVGPRTYNTLSYIGSGAYGMVCSAY 43
QY 61 DHVTRKTRVAIKKISPEFHQYQCORTLREIQILLRFRHENVIGRIDLRASTLEMRDYYI 120
DB 44 DVNKKVRVAIKKISPEFHQYQCORTLREIKILLRFRHENIIGINDIIRAPTIEQMKVYI 103
QY 121 VQDMETDLYKLLKSKQSLNDHYCYFLYQILRGKXTHSANVLRDLKPSNLLINTTCDL 180
DB 104 VQDMETDLYKLLKTOHLSNDHYCYFLYQILRGKXTHSANVLRDLKPSNLLINTTCDL 163
QY 181 KICDFGLARADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 164 KICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 223
QY 241 NRPIFGKHYLDQNLHILGILGSPSOEDLNCIINMKARNYLOSLPSKTKVAKLFPKSD 300
DB 224 NRPIFGKHYLDQNLHILGILGSPSOEDLNCIINMKARNYLLSLPHKKNKVPWNRLPFNAD 283
QY 301 SKALDLDRLMTENPKRITV-----AEPPTFAMELDDLPKRL 340
DB 284 SKALDLDRLMTFNPHKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMLDLPKRL 343
QY 341 KELIFOETARFQPG 354
DB 344 KELIFETARFQPG 357
RESULT 14
Q8UWG6 CHICK
ID Q8UWG6 CHICK PRELIMINARY; PRT; 368 AA.
AC Q8UWG6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-regulated kinase 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC	Gallus.	DT	13-SEP-2005 (Rel. 48, Last annotation update)
OX	NCBI_TaxID=9031;	DE	Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
RN	[1]	DE	signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
RN	NUCLEOTIDE SEQUENCE.	DE	2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERT1).
RC	TISSUE=Pineal gland;	GN	Name=Mapki; Synonyms=Erk2, Mapk, Prkml;
RA	Kasahara T., Higashi C., Okano T., Fukada Y.;	OS	Name=musculus (Mouse).
RA	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	- - CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.	OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
CC	- - SIMILARITY: Belongs to the Ser/Thr protein kinase family.	OC	Muroidea; Muridae; Murinae; Mus.
DR	EMBL; AY033635; AAK56503.1; -; mRNA.	OX	NCBI_TaxID=10090;
DR	HSSP; P28482; IPME.	RN	[1]
DR	SMR; Q8UWG6; 24-367.	RP	NUCLEOTIDE SEQUENCE.
DR	Ensembl; ENSGALG00000004501; Gallus gallus.	RC	STRAIN=Swiss; TISSUE=Fibroblast;
DR	GO; GO:003524; F:ATP binding; IEA.	RX	MEDLINE=91305126; PubMed=1649458;
DR	GO; GO:0004707; F:MAP kinase activity; IEA.	RA	Her J.-H., Wu J.-S., Rall T.B., Sturgill T.W., Weber M.J.;
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.	RT	"Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.	RT	tyrosine phosphorylation."
DR	InterPro; IPR008349; Erk 1.2 MAPK.	RL	Nucleic Acids Res. 19:3743-3743(1991).
DR	InterPro; IPR003527; MAP_kin.	RN	[2]
DR	InterPro; IPR000719; Prot_kinase.	RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
DR	InterPro; IPR008271; Ser_Thr_kin_AS.	RC	STRAIN=C57BL/6J, and NOD; TISSUE=Head, Thymus, and Urinary bladder;
DR	Pfam; PF00069; Pkinase; 1.	RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
DR	PRINTS; PR01770; ERK1ERK2MAPK.	RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
DR	ProDom; PD000001; Prot_kinase; 1.	RA	Nikaido I., Otsu N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
DR	SMART; SM00220; S_TKc_1.	RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
DR	PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.	RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	RA	Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
KW	ATP-binding; Kinase; Nucleotide-binding;	RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
KW	Serine/threonine-protein kinase; Transferase.	RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
SQ	SEQUENCE 368 AA; 41942 MW; A6F3A3D27C6C9A44 CRC64;	RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
		RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
		RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
		RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
		RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
		RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
		RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
		RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
		RA	Sallana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
		RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
		RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
		RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
		RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
		RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
		RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
		RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
		RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
		RA	Birney E., Hayashizaki Y.;
		RT	"Analysis of the mouse transcriptome based on functional annotation of
		RT	60,770 full-length cDNAs."
		RL	Nature 420:563-573(2002).
		RN	[3]
		RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
		RC	TISSUE=Eye;
		RX	MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
		RA	Stausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
		RA	Klauser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
		RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
		RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
		RA	Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
		RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
		RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
		RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
		RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
		RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
		RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
		RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
		RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
		RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
		RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
		RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
		RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.,
		RT	"Generation and initial analysis of more than 15,000 full-length human

Query Match	81.8%; Score 1537; DB 2; Length 368;
Best Local Similarity	78.9%; Pred. No. 1.2e-89;
Matches	295; Conservative 19; Mismatches 28; Indels 32; Gaps 2;
QY	1 MAAAAAGGGGPRTEGPGVPEVEMVKQPEVDGPRYTOLOYIGEGAYGMVSSAY 60
DB	4 VAGAAAAGGSAG-----AGGPEVVRGQVDFGPRYTNLSYIGEGAYGMVCSAY 51
QY	61 DHVRKTRVAIKLSPFEHOTYCORTURETOILLRPHENVIGRIDTLRASTLEAMRDVYI 120
DB	52 DNNKVRVAIKLSPFEHOTYCORTUREIKLLRPHENIIGINDIIRAPTEQMDVYI 111
QY	121 VQDLMETDLYKLLKSQSLNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB	112 VQDLMETDLYKLLKTQHLNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 171
QY	181 KICDFGLARIADPEHDTGTLTYVATRVYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
DB	172 KICDFGLARVADPDHDTGTLTYVATRVYRAPEIMLSKGYTKSDIWSVGCILAEMLS 231
QY	241 NRPIFPKHYLDNLHILGLSPSODLNCIIMKARNVLOSPLKTKVNAKLPKSD 300
DB	232 NRPIFPKHYLDNLHILGLSPSODLNCIIMKARNVLOSPLKTKVNAKLPKSD 291
QY	301 SKALDLDLRLMTFNPKNRI-----TVAEPTTFAMELDDLPKRL 340
DB	292 PKALDLDLRLMTFNPKNRI-----TVAEPTTFAMELDDLPKRL 351
QY	341 KELIFQETARFQSG 354
DB	352 KELIFQETARFQSG 365
RESULT 15	
MOUSE	
ID	MK01_MOUSE
AC	P63085; P27703;
DT	01-AUG-1992 (Rel. 23, Last sequence update)
DT	01-AUG-1992 (Rel. 23, Last sequence update)

and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
[4]
NUCLEOTIDE SEQUENCE OF 151-189.
RC STRAIN=CBA; TISSUE=Bone marrow;
RX MEDLINE=93185941; PubMed=8444355; DOI=10.1016/0378-1119(93)90411-U;
RA Ershler M.A.; Nagorskaya T.V.; Visser J.W.M.; Belyavsky A.V.;
RT "Novel CDC2-related protein kinases produced in murine hematopoietic
stem cells.";
RL Gene 124:305-306(1993).
[5]
PROSPHORYLATION SITES THR-183 AND TYR-185, AND PARTIAL PROTEIN
SEQUENCE.
RP MEDLINE=91184134; PubMed=1849075;
RA Payne D.M., Rossomando A.J., Martino P., Erickson A.K., Her J.-H.,
RA Shabanowitz J., Hunt D.F., Weber M.J., Sturgill T.W.;
RT "Identification of the regulatory phosphorylation sites in
PP42/mitogen-activated protein kinase (MAP kinase).";
RL EMBO J. 10:885-892(1991).
CC -1- FUNCTION: Involved in both the initiation and regulation of
meiosis, mitosis, and postmitotic functions in differentiated
cells by phosphorylating a number of transcription factors such as
CC EXK-1. Phosphorylates E14EBP1; required for initiation of
CC translation. Phosphorylates microtubule-associated protein 2
(MAP2) (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
threonine in response to insulin and NGF. Both phosphorylations
are required for activity (By similarity).
CC -1- TISSUE SPECIFICITY: Widely expressed.
CC -1- PFM: Autophosphorylated on threonine and tyrosine residues, which
correlates with a slow and low level of autoactivation (By
similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
kinase subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; X58712; CAA41548.1; -; mRNA.
EMBL; AK035386; BAC29053.1; -; mRNA.
EMBL; AK048127; BAC33251.1; -; mRNA.
EMBL; AK087925; BAC40044.1; -; mRNA.
EMBL; BC058258; AAH58258.1; -; mRNA.
EMBL; D10939; BAA01733.1; -; mRNA.
PIR; S16444; S16444.
SMR; P63085; 14-357.
IntAct; P63085; -.
MGI; MGI:1346858; Mapk1.
GO; GO:0005737; C:cytoplasm; IDA.
GO; GO:0005634; C:nucleus; IDA.
GO; GO:0004707; F:MAP kinase activity; IDA.
GO; GO:0001784; F:phosphotyrosine binding; IMP.
GO; GO:0005515; F:protein binding; IPI.
GO; GO:0004672; F:protein kinase activity; IDA.
GO; GO:0000165; P:MAPKK cascade; IDA.
GO; GO:0009887; P:organogenesis; IDA.
GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
GO; GO:0006974; P:response to DNA damage stimulus; IDA.
GO:GO:0007165; P:signal transduction; TAS.
InterPro; IPR008349; Erk 1.2 MAPK.
InterPro; IPR003527; MAP kin.
InterPro; IPR000719; Prot Kinase.
InterPro; IPR008271; Ser Thr_pkin_AS.
Efam; EF00069; Pkinase; I.
PRINTS; PR01770; ERKIERK2MAPK.
PROSITE; PS01351; MAPK; 1.
PROSITE; PS00107; PROTEIN KINASE ATP; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Cell cycle; Direct protein sequencing; Kinase;
KW Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase;
KW Transferase.
FT DOMAIN 23 311 Protein kinase.
FT NP_BIND 29 37 ATP (By similarity).
FT COMBIAS 2 7 Poly-Ala.
FT ACT_SITE 147 147 By similarity.
FT BINDING 52 52 ATP (By similarity).
FT MOD_RES 183 183 Phosphothreonine (activates the kinase).
FT MOD_RES 185 185 Phosphotyrosine (activates the kinase).
SQ SEQUENCE 358 AA; 41276 MW; 3BBCF22471EDBA0B CRC64;

Query Match 81.3%; Score 1528.5; DB 1; Length 358;
Best Local Similarity 78.8%; Pred. No. 4.2e-89;
Matches 294; Conservative 19; Mismatches 22; Indels 39; Gaps 2;

QY 1 MAAAAAGGGGGPRRTGPGVGPGEVEMVKGQPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAG-----PENVRGQVDFVGPRTNLSYIGEGAYGMVCSAY 41

QY 61 DHVTRKTRVAIKKISPFHQTYCORTLREIQILLRFHENVIGIRDLRASTLEAMRDVYI 120
DB DNLNKKVRVAIKKISPFHQTYCORTLREIKILLRFHENVIGIRDLRASTLEAMRDVYI 101

QY 121 VQDLMETDLYKLKLSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLINTTCDL 180
DB VQDLMETDLYKLKLSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLINTTCDL 161

QY 181 KICDFGLARADPEHDHDTGFLTEVATRYTRAPEIMLSNGYTKSIDINSGVGCILAEMLS 240
DB KICDFGLARVADPDHDTGFLTEVATRYTRAPEIMLSNGYTKSIDINSGVGCILAEMLS 221

QY 241 NRPIFFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLOSLPSKTKVAKLPKSD 300
DB NRPIFFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLOSLPSKTKVAKLPKSD 281

QY 301 SKALDLLDRMLTNPKNRITY-----ASEPTFAMELDDLPKRL 340
DB SKALDLLDRMLTNPKNRITY-----ASEPTFAMELDDLPKRL 341

QY 341 KELIFORTARFPQG 354
DB KELIFORTARFPQG 355

Search completed: February 6, 2006, 15:34:39
Job time : 172.223 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:34:56 ; Search time 36.2104 Seconds
(without alignments)
819.670 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAGGGGGPRRTGV.....LKELIFQETARFQGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1859	98.9	379	2	US-08-622-277A-8
2	1859	98.9	379	2	US-09-411-628-14
3	1859	98.9	379	2	US-10-174-794-14
4	1859	98.9	379	2	US-09-538-092-1021
5	1853	98.6	379	2	US-09-025-580-25
6	1853	98.6	379	2	US-09-642-749-25
7	1853	98.6	631	2	US-09-417-197-39
8	1787	95.1	624	2	US-09-417-197-57
9	1735	92.3	367	1	US-08-176-620A-11
10	1735	92.3	367	1	US-08-461-985-11
11	1734	92.3	355	2	US-08-459-953A-8
12	1734	92.3	355	2	US-09-393-212-8
13	1723	91.7	365	1	US-08-176-620A-2
14	1723	91.7	365	1	US-08-463-862-2
15	1723	91.7	365	1	US-08-461-985-2
16	1723	91.7	365	1	US-08-458-887-2
17	1723	91.7	365	2	US-08-932-787B-2
18	1723	91.7	365	2	US-08-932-012C-2
19	1723	91.7	365	2	US-08-888-818C-2
20	1553.5	82.7	380	2	US-09-949-016-9919
21	1548.5	82.4	380	2	US-08-459-953A-9
22	1548.5	82.4	380	2	US-09-393-212-9
23	1538.5	81.9	360	2	US-09-457-040B-3
24	1538.5	81.9	360	2	US-09-538-092-1028
25	1530	81.4	415	1	US-08-176-620A-4
26	1530	81.4	415	1	US-08-463-862-4
27	1530	81.4	415	1	US-08-461-985-4

28	1530	81.4	415	1	US-08-458-887-4
29	1530	81.4	415	2	US-08-932-787B-4
30	1530	81.4	415	2	US-08-932-012C-4
31	1530	81.4	415	2	US-08-888-818C-4
32	1529.5	81.4	360	2	US-08-622-277A-14
33	1528.5	81.3	358	1	US-08-176-620A-12
34	1528.5	81.3	358	1	US-08-461-985-12
35	1528.5	81.3	358	2	US-09-457-040B-4
36	1528.5	81.3	364	2	US-09-417-197-59
37	1528.5	81.3	604	2	US-09-417-197-41
38	1528.5	81.3	605	2	US-09-457-040B-35
39	1515.5	80.7	365	2	US-09-025-580-2
40	1509.5	80.3	360	2	US-09-642-749-2
41	1509.5	80.3	360	2	US-08-176-620A-8
42	905	48.2	183	1	US-08-461-985-8
43	905	48.2	183	2	US-08-932-787B-16
44	905	48.2	183	2	US-08-932-012C-16
45	905	48.2	183	2	US-08-932-012C-16

ALIGNMENTS

RESULT 1
US-08-622-277A-8
; Sequence 8, Application US/08622277A
; Patent No. 6001580
; GENERAL INFORMATION:
; APPLICANT: Tani, Akiyoshi
; APPLICANT: Ichimori, Yuzo
; TITLE OF INVENTION: Method For Assaying MAP Kinase
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,277A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 320577-1995
; FILING DATE: 08-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 305456-1995
; FILING DATE: 24-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 070125-1995
; FILING DATE: 28-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toohy, Kimberlin M.
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 04221.0039-00000
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
US-08-622-277A-8

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRTEGPGVPGVEVMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGPGVPGVEVMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240

QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300

QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPPPTFAMELDDLPKERL 340
DB 301 SKALDLDRLMTFNPKNKRIT-----VAEPPPTFAMELDDLPKERL 340

QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 2
US-09-411-628-14
; Sequence 14, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-411-628-14

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRTEGPGVPGVEVMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGPGVPGVEVMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240

DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPPPTFAMELDDLPKERL 340
DB 301 SKALDLDRLMTFNPKNKRITVEEALAHPLYEQYDPTDEPVAEPPPTFAMELDDLPKERL 360
QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 3
US-10-174-794-14
; Sequence 14, Application US/10174794
; Patent No. 6664086
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-174-794-14

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRTEGPGVPGVEVMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGPGVPGVEVMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240

QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPPPTFAMELDDLPKERL 340
DB 301 SKALDLDRLMTFNPKNKRITVEEALAHPLYEQYDPTDEPVAEPPPTFAMELDDLPKERL 360
QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 4
US-09-538-092-1021

Sequence 1021, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Giot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: 60/127,352
PRIORITY FILING DATE: 1999-04-01
PRIORITY APPLICATION NUMBER: 60/178,965
PRIORITY FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuraPatSeqFormat Version 0.9
SEQ ID NO 1021
LENGTH: 379
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)-(0)
OTHER INFORMATION: Polypeptide Accession Number P27361
US-09-538-092-1021

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60

Qy 61 DHVKTVAIKTISPFHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVKTVAIKTISPFHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

Qy 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180

Qy 181 KICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240

Qy 241 NRPIFGKHLYDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Db 241 NRPIFGKHLYDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300

Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDRLMTFNPKNRITVEALAHPLYEQYYDPTDEPVAEPPTFAMELDDLPKERL 360

Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 5
US-09-025-580-25
Sequence 25, Application US/09025580
Patent No. 6162613
GENERAL INFORMATION:
APPLICANT: Su, Michael Shin-San
APPLICANT: Fox, Ted
APPLICANT: Wilson, Keith Phillip
APPLICANT: Germann, Ursula A.
TITLE OF INVENTION: Methods For Designing Inhibitors of
TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas

CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,580
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haley, James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: VPI 97-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-580-25

Query Match 98.6%; Score 1853; DB 2; Length 379;
Best Local Similarity 94.5%; Pred. No. 5.2e-185;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

Qy 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60

Qy 61 DHVKTVAIKTISPFHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVKTVAIKTISPFHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

Qy 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180

Qy 181 KICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240

Qy 241 NRPIFGKHLYDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Db 241 NRPIFGKHLYDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300

Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDRLMTFNPKNRITVEALAHPLYEQYYDPTDEPVAEPPTFAMELDDLPKERL 360

Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 6
US-09-642-749-25
Sequence 25, Application US/09642749
Patent No. 6849716
GENERAL INFORMATION:
APPLICANT: Su, Michael Shin-San
APPLICANT: Fox, Ted
APPLICANT: Wilson, Keith Phillip
APPLICANT: Germann, Ursula A.
TITLE OF INVENTION: Methods For Designing Inhibitors of
TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase

; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/642,749
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-642-749-25

Query Match 98.6%; Score 1853; DB 2; Length 379;
Best Local Similarity 94.5%; Pred. No. 5.2e-185;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRRTGPGVPGVEVMVGQPPDVGPRTYQLQYIGEGAYGMVSSAY 60
Db 1 MAAAAQGGGGPRRTGPGVPGVEVMVGQPPDVGPRTYQLQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLPFKSD 300
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLPFKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKERL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 360
QY 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 7
US-09-417-197-39
; Sequence 39, Application US/09417197
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MERK1-F64L-S65T-GFP fusion
US-09-417-197-57
Query Match 95.1%; Score 1787; DB 2; Length 624;

; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-09-417-197-39

Query Match 98.6%; Score 1853; DB 2; Length 631;
Best Local Similarity 94.5%; Pred. No. 1.1e-184;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRRTGPGVPGVEVMVGQPPDVGPRTYQLQYIGEGAYGMVSSAY 60
Db 253 MAAAAQGGGGPRRTGPGVPGVEVMVGQPPDVGPRTYQLQYIGEGAYGMVSSAY 312
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 313 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 372
QY 121 VQDLMETDLYKLLKSQQLSNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 373 VQDLMETDLYKLLKSQQLSNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 432
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
Db 433 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 492
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLPFKSD 300
Db 493 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAVAKLPFKSD 552
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKERL 340
Db 553 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 612
QY 341 KELIFQETARFQGVLEAP 359
Db 613 KELIFQETARFQGVLEAP 631

RESULT 8
US-09-417-197-57
; Sequence 57, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MERK1-F64L-S65T-GFP fusion
US-09-417-197-57
Query Match 95.1%; Score 1787; DB 2; Length 624;

```
;
; MOLECULE TYPE: protein
; US-08-176-620A-11
;
; TOPOLOGY: unknown
;
; Query Match 92.3%; Score 1735; DB 1; Length 367;
; Best Local Similarity 91.8%; Pred. No. 1.1e-172;
; Matches 347; Conservative 1; Mismatches 9; Indels 20; Gaps 1;
;
; QY 2 AAAAAAGGGGGRRRTGGVPGVGEVEMVKGQPFVGPRTYQIYIGEGAYGMVSSAYD 61
; DB 3 AAAAAAGGGGGRRRTAGVVPVGEVEMVKGQPFVGPRTYQIYIGEGAYGMVSSAYD 62
; QY 62 HVKTRVAIKKISPFHEQTYCORTLEIQLLFRHENVIGIRDLRASTLEAMRDVYIV 121
; DB 63 HVKTRVAIKKISPFHEQTYCORTLEIQLLFRHENVIGIRDLRASTLEAMRDVYIV 122
; QY 122 QDLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLK 181
; DB 123 QDLMETDLYKLLKSKQSLNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLK 182
; QY 182 ICDPGLARIADPHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 241
; DB 183 ICDPGLARIADPHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 242
; QY 242 RPIFPKGHYLDQNLHILGILGSPQEDLNCIIINMKARNYLQSLPSKTKVAVAKLPPKSDS 301
; DB 243 RPIFPKGHYLDQNLHILGILGSPQEDLNCIIINMKARNYLQSLPSKTKVAVAKLPPKSDS 302
; QY 302 KALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLKPKERLK 341
; DB 303 KALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFDMELDDLKPKERLK 362
; QY 342 ELIFQETARFQPGVLEAP 359
; DB 363 ELIFQETARFQGAPEGP 380
;
; RESULT 9
; US-08-176-620A-11
; Sequence 11, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,620A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
;
; QY 13 EPRRTGGVPGVGEVEMVKGQPFVGPRTYQIYIGEGAYGMVSSAYDHRKTRVAIKK 72
; DB 1 EPRGTAGVVPVPGVGEVEMVKGQPFVGPRTYQIYIGEGAYGMVSSAYDHRKTRVAIKK 60
; QY 73 ISPFHEQTYCORTLEIQLLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKL 132
; DB 61 ISPFHEQTYCORTLEIQLLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKL 120
; QY 133 LKSKQSLNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIAD 192
; DB 121 LKSKQSLNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIAD 180
; QY 193 PEHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYD 252
; DB 181 PEHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYD 240
; QY 253 QLNHLILGILGSPQEDLNCIIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLDRLMT 312
; DB 241 QLNHLILGILGSPQEDLNCIIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLDRLMT 300
; QY 313 FNPKNRIT-----VAEPPTFAMELDDLKPKERLKELIFQETARFQ 352
; DB 301 FNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFDMELDDLKPKERLKELIFQETARFQ 360
; QY 353 PGVLEAP 359
; DB 361 PGAPAP 367
;
; RESULT 10
; US-08-461-985-11
; Sequence 11, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-461-985-11

Query Match          92.3%; Score 1735; DB 1; Length 367;
Best Local Similarity 91.8%; Pred. No. 1.1e-172;
Matches 337; Conservative 1; Mismatches 9; Indels 20; Gaps 1;

QY 13 EPRTEGVGPGFGEVEMVKQPFVGPRTQYIGEGAYGMVSSAYDHRVTRVAIKK 72
Db 1 EPRGTAGVVPVPGVEVVKQPFVGPRTQYIGEGAYGMVSSAYDHRVTRVAIKK 60
QY 73 ISPFHQTCQRTLRIOILLRPHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLYKL 132
Db 61 ISPFHQTCQRTLRIOILLRPHENVIGIRDLIRAPTLEAMRDVYIVQDLMETDLYKL 120
QY 133 LKSQLSNDHICVFLYQILRGLKVIHSANVLHRDLKPSNLLINTTCDLKI CDFGLARIAD 192
Db 121 LKSQLSNDHICVFLYQILRGLKVIHSANVLHRDLKPSNLLINTTCDLKI CDFGLARIAD 180
QY 193 PEHDHTGFLTEYVATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPFPGRKHYLD 252
Db 181 PEHDHTGFLTEYVATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPFPGRKHYLD 240
QY 253 QLNHIILGILGPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSDSKALDLDRLMT 312
Db 241 QLNHIILGILGPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSDSKALDLDRLMT 300
QY 313 FPNKRIT-----VAEPPFTFAMELDLDPKERLKELIFOETARFQVLEAP 352
Db 301 FPNKRITVEEALAHPLYEQYDPTDFVAEPPFTFAMELDLDPKERLKELIFOETARFQ 360
QY 353 PGVLEAP 359
Db 361 PGVLEAP 367

RESULT 11
US-08-459-953A-8
; Sequence 8, Application US/08459953A
; Patent No. 6030822
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,953A
; FILING DATE: June 2, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,494
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-953A-8

Query Match          92.3%; Score 1734; DB 2; Length 355;
Best Local Similarity 94.4%; Pred. No. 1.3e-172;
Matches 335; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 25 PGEVEMVKQPFVGPRTQYIGEGAYGMVSSAYDHRVTRVAIKKISPEHOTYCOR 84
Db 1 PGEVEMVKQPFVGPRTQYIGEGAYGMVSSAYDHRVTRVAIKKISPEHOTYCOR 60
QY 85 TUREIQILRPHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC 144
Db 61 TUREIQILRPHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC 120
QY 145 YELYQILRGLKVIHSANVLHRDLKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLTEY 204
Db 121 YELYQILRGLKVIHSANVLHRDLKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLTEY 180
QY 205 VATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPFPGRKHYLDQNLHILGLSP 264
Db 181 VATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPFPGRKHYLDQNLHILGLSP 240
QY 265 SQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRIT- 320
Db 241 SQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRITVEEA 300
QY 321 -----VAEPPFTFAMELDLDPKERLKELIFOETARFQVLEAP 359
Db 301 LAHPYLEQYDPTDFVAEPPFTFAMELDLDPKERLKELIFOETARFQVLEAP 355

RESULT 12
US-09-393-212-8
; Sequence 8, Application US/09393212
; Patent No. 6579972
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/393.212
FILING DATE: 09-Sep-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/459,953A
FILING DATE: June 2, 1995
APPLICATION NUMBER: 08/029,494
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-393-212-8

Query Match 92.3%; Score 1734; DB 2; Length 355;
Best Local Similarity 94.4%; Pred. No. 1.3e-172;
Matches 335; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 25 PGEVVMKGPFDVGPRTYQIQYIGEGAYGMVSSAYDHRVTRVAIKKISPFHEHTYQCR 84
DB 1 PGEVVMKGPFDVGPRTYQIQYIGEGAYGMVSSAYDHRVTRVAIKKISPFHEHTYQCR 60
QY 85 TLREIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLKSQLSNDHIC 144
DB 61 TLREIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLKSQLSNDHIC 120
QY 145 YFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPEHDTGFLTEY 204
DB 121 YFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPEHDTGFLTEY 180
QY 205 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQLNHILGILGSP 264
DB 181 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQLNHILGILGSP 240
QY 265 SQEDLNCIINNKARNYLSQSPSKTKVAKLFPKSDSKALDLDLDRMLTFNPNKRIT--- 320
DB 241 SQEDLNCIINNKARNYLSQSPSKTKVAKLFPKSDSKALDLDLDRMLTFNPNKRITVERA 300
QY 321 -----VAEEPTFAMELDLDPKRLKELIFQETARFPQGVLEAP 359
DB 301 LAHPYLEQYDPTDEPVAEEPTFAMELDLDPKRLKELIFQETARFPQGVLEAP 355

RESULT 13
US-08-176-620A-2
; Sequence 2, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Fanayotatos, Nikos
; TITLE OF INVENTION: A family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,620A
FILING DATE: 03-JAN-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-620A-2

Query Match 91.7%; Score 1723; DB 1; Length 365;
Best Local Similarity 91.8%; Pred. No. 1.9e-171;
Matches 335; Conservative 1; Mismatches 9; Indels 20; Gaps 1;
QY 15 RRTGEGVGPVGEVVMKGPFDVGPRTYQIQYIGEGAYGMVSSAYDHRVTRVAIKKIS 74
DB 1 RGTAGVVPVPGVGEVVMKGPFDVGPRTYQIQYIGEGAYGMVSSAYDHRVTRVAIKKIS 60
QY 75 PPEHTYQVQRTLREIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLK 134
DB 61 PPEHTYQVQRTLREIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLK 120
QY 135 SQLSNDHICYFYQILRLGKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE 194
DB 121 SQLSNDHICYFYQILRLGKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE 180
QY 195 HDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQL 254
DB 181 HDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQL 240
QY 255 NHILGILGSPSQEDLNCIINNKARNYLSQSPSKTKVAKLFPKSDSKALDLDLDRMLTFN 314
DB 241 NHILGILGSPSQEDLNCIINNKARNYLSQSPSKTKVAKLFPKSDSKALDLDLDRMLTFN 300
QY 315 PNKRIT-----VAEEPTFAMELDLDPKRLKELIFQETARFPQGVLEAP 354
DB 301 PNKRITVERALAHPLYEQYDPTDEPVAEEPTFAMELDLDPKRLKELIFQETARFPQGV 360
QY 355 VLEAP 359
DB 361 APEAP 365

RESULT 14
US-08-463-862-2
; Sequence 2, Application US/08463862
; Patent No. 5776751
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G. et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463,862
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/701,544
 FILING DATE: 16-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6526-049
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 365 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-862-2

Query Match 91.7%; Score 1723; DB 1; Length 365;
 Best Local Similarity 91.8%; Pred. No. 1.9e-171;
 Matches 335; Conservative 1; Mismatches 9; Indels 20; Gaps 1;
 QY 15 RTEGVGPGVGEVEMVKGQPFVGPRTQLOVIGEGAYGMVSSAYDHYRKTVAIKKIS 74
 DB 1 RGTAGVVPVPGVEVVKQPFVGPRTQLOVIGEGAYGMVSSAYDHYRKTVAIKKIS 60
 QY 75 PFEHOTYQCTREIQILRLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK 134
 DB 61 PFEHOTYQCTREIQILRLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK 120
 QY 135 SQQLSNDHICFYLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE 194
 DB 121 SQQLSNDHICFYLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE 180
 QY 195 HDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIVSVCILAEMLSNRPIFPKGHYLDOL 254
 DB 181 HDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIVSVCILAEMLSNRPIFPKGHYLDOL 240
 QY 255 NHILGILGSPSQEDLNCIINMKARNYIQLSPKTKVAVAKLFPKSDSKALDLDRLMTFN 314
 DB 241 NHILGILGSPSQEDLNCIINMKARNYIQLSPKTKVAVAKLFPKSDSKALDLDRLMTFN 300
 QY 315 PNKRIT-----VAEPFTFAMELDDLPKRLKELIFQETARFQPG 354
 DB 301 PNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFDMELDDLPKRLKELIFQETARFQPG 360
 QY 355 VLEAP 359
 DB 361 APEAP 365

RESULT 15

US-08-461-985-2
 ; Sequence 2, Application US/08461985
 ; Patent No. 5872006
 ; GENERAL INFORMATION:
 ; APPLICANT: Boulton, Teri G.

APPLICANT: Cobb, Melanie H.
 APPLICANT: Yancopoulos, George D.
 APPLICANT: Nye, Steven
 APPLICANT: Panayotatos, Nikos
 TITLE OF INVENTION: A Family of Map2 Protein Kinases
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/461,985
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/176,620
 FILING DATE: 03-JAN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6526-123
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 365 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-461-985-2

Query Match 91.7%; Score 1723; DB 1; Length 365;
 Best Local Similarity 91.8%; Pred. No. 1.9e-171;
 Matches 335; Conservative 1; Mismatches 9; Indels 20; Gaps 1;
 QY 15 RTEGVGPGVGEVEMVKGQPFVGPRTQLOVIGEGAYGMVSSAYDHYRKTVAIKKIS 74
 DB 1 RGTAGVVPVPGVEVVKQPFVGPRTQLOVIGEGAYGMVSSAYDHYRKTVAIKKIS 60
 QY 75 PFEHOTYQCTREIQILRLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK 134
 DB 61 PFEHOTYQCTREIQILRLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK 120
 QY 135 SQQLSNDHICFYLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE 194
 DB 121 SQQLSNDHICFYLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE 180
 QY 195 HDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIVSVCILAEMLSNRPIFPKGHYLDOL 254
 DB 181 HDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIVSVCILAEMLSNRPIFPKGHYLDOL 240
 QY 255 NHILGILGSPSQEDLNCIINMKARNYIQLSPKTKVAVAKLFPKSDSKALDLDRLMTFN 314
 DB 241 NHILGILGSPSQEDLNCIINMKARNYIQLSPKTKVAVAKLFPKSDSKALDLDRLMTFN 300
 QY 315 PNKRIT-----VAEPFTFAMELDDLPKRLKELIFQETARFQPG 354
 DB 301 PNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFDMELDDLPKRLKELIFQETARFQPG 360
 QY 355 VLEAP 359
 DB 361 APEAP 365

Search completed: February 6, 2006, 15:36:49
Job time : 37.2104 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:35:46 ; Search time 125.184 Seconds
(without alignments)
1198.238 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAGGGGGGPRTEGV.....LKELIFQETARFQGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications_AA_Main:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pgp:*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pgp:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pgp:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pgp:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pgp:*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pgp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1879	100.0	359	5	US-10-623-108-4
2	1879	100.0	359	5	US-10-623-108-8
3	1873	99.7	359	4	US-10-114-270-110
4	1859	98.9	379	4	US-10-233-448-7
5	1859	98.9	379	4	US-10-174-794-14
6	1859	98.9	379	4	US-10-394-322A-19
7	1859	98.9	379	4	US-10-701-490-8
8	1859	98.9	379	5	US-10-735-118-3
9	1859	98.9	379	5	US-10-623-108-6
10	1859	98.9	379	5	US-10-945-684-2
11	1853	98.6	379	4	US-10-060-065-16
12	1853	98.6	379	4	US-10-059-585-37
13	1853	98.6	379	4	US-10-170-663-25
14	1853	98.6	379	5	US-10-893-072-25
15	1853	98.6	631	4	US-10-072-036-39
16	1787	95.1	624	4	US-10-072-036-57
17	1734	92.3	355	4	US-10-461-402-8
18	1599	85.1	335	5	US-10-623-108-2
19	1550.5	82.5	392	5	US-10-732-923-1420
20	1548.5	82.4	380	4	US-10-461-402-9
21	1538.5	81.9	360	4	US-10-171-311-127
22	1538.5	81.9	360	4	US-10-233-448-8
23	1538.5	81.9	360	4	US-10-301-822-115
24	1538.5	81.9	360	4	US-10-394-322A-20
25	1538.5	81.9	360	5	US-10-941-635-56
26	1532.5	81.6	360	4	US-10-664-421-56
27	1528.5	81.3	358	5	US-10-945-684-1

28	1528.5	81.3	358	5	US-10-451-291B-10	Sequence 10, Appl
29	1528.5	81.3	604	4	US-10-072-036-59	Sequence 59, Appl
30	1528.5	81.3	605	4	US-10-072-036-41	Sequence 41, Appl
31	1523	81.1	356	4	US-10-664-421-19	Sequence 19, Appl
32	1520	80.9	355	5	US-10-941-635-19	Sequence 19, Appl
33	1518	80.8	348	5	US-10-756-149-4856	Sequence 4856, Ap
34	1509.5	80.3	360	4	US-10-170-663-2	Sequence 2, Appl1
35	1509.5	80.3	360	5	US-10-893-072-2	Sequence 2, Appl1
36	1508	80.3	369	5	US-10-732-923-1421	Sequence 1421, Ap
37	1445	76.9	351	5	US-10-732-923-1406	Sequence 1406, Ap
38	1418	75.5	369	5	US-10-732-923-1456	Sequence 1456, Ap
39	1384.5	73.7	395	5	US-10-732-923-1424	Sequence 1424, Ap
40	1371.5	73.0	362	5	US-10-732-923-1418	Sequence 1418, Ap
41	1317	70.1	376	4	US-10-369-493-5689	Sequence 5689, Ap
42	944	50.2	352	5	US-10-732-923-1357	Sequence 1357, Ap
43	944	50.2	352	5	US-10-732-923-1367	Sequence 1367, Ap
44	941	50.1	355	5	US-10-732-923-1495	Sequence 1495, Ap
45	940	50.0	353	5	US-10-732-923-1492	Sequence 1492, Ap

ALIGNMENTS

RESULT 1

US-10-623-108-4
; Sequence 4, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-4

Query Match 100.0%; Score 1879; DB 5; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.3e-147;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAAAAGGGGGPRTEGVGPGVGEVENVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60
Db	1	MAAAAAGGGGGPRTEGVGPGVGEVENVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60
Qy	61	DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI	120
Db	61	DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI	120
Qy	121	VQDLMETDLYKLKLSQOLSNHICYFLYQTLRLGKVIHNSANVLRDLKPSNLINTTCDL	180
Db	121	VQDLMETDLYKLKLSQOLSNHICYFLYQTLRLGKVIHNSANVLRDLKPSNLINTTCDL	180
Qy	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240
Db	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240
Qy	241	NRPIFFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLOSLPKTKVAKLPPKSD	300
Db	241	NRPIFFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLOSLPKTKVAKLPPKSD	300
Qy	301	SKALDLDRLMTFNPKNRTVAEPEPTFAMELDDPKERLKELIFOETARFQGVLEAP	359
Db	301	SKALDLDRLMTFNPKNRTVAEPEPTFAMELDDPKERLKELIFOETARFQGVLEAP	359

RESULT 2

US-10-623-108-8

; Sequence 8, Application US/10623108

```
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAL, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-8

Query Match      100.0%; Score 1879; DB 5; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.3e-147;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAAQGGGGEPRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAQGGGGEPRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
   |||||

QY 61 DHVKTTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVKTTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||

QY 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||
Db 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
   |||||

QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||||
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||||

QY 301 SKALDLLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||||
Db 301 SKALDLLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||||

RESULT 3
US-10-114-270-110
; Sequence 110, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangoli, Ssha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
```

```
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: Macdougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 110
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-110

Query Match      99.7%; Score 1873; DB 4; Length 359;
Best Local Similarity 99.7%; Pred. No. 1.4e-146;
Matches 358; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAAAQGGGGEPRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAQGGGGEPRTGEGVGPGEVEMVKGQPDFVGPRTQLOYGEGAYGMVSSAY 60
   |||||

QY 61 DHVKTTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVKTTRVAIKKISPEHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||

QY 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||
Db 121 VDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
   |||||

QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||||
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||||

QY 301 SKALDLLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||||
Db 301 SKALDLLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||||

RESULT 4
US-10-233-448-7
; Sequence 7, Application US/10233448
; Publication No. US20030109419A1
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
```

APPLICANT: Fisone, Gilberto
; TITLE OF INVENTION: METHOD FOR CLASSIFICATION OF ANTI-PSYCHOTIC DRUGS
; FILE REFERENCE: 11181-009
; CURRENT APPLICATION NUMBER: US/10/233,448
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 60/316,338
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-448-7

Query Match 98.9%; Score 1859; DB 4; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPPDVGPRTYQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPPDVGPRTYQLOVIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEFHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEFHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAAWAKLFPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAAWAKLFPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKERL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVAEPEPTFAMELDDLPKERL 360
QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 5
US-10-174-794-14
; Sequence 14, Application US/10174794
; Publication No. US20030166220A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-174-794-14

Query Match 98.9%; Score 1859; DB 4; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;

MATCHES 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPPDVGPRTYQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPPDVGPRTYQLOVIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEFHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEFHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAAWAKLFPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAAWAKLFPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKERL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVAEPEPTFAMELDDLPKERL 360
QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 6
US-10-394-322A-19
; Sequence 19, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-19

Query Match 98.9%; Score 1859; DB 4; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPPDVGPRTYQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPPDVGPRTYQLOVIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEFHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEFHQTYCQRTLRREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAAWAKLFPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAAWAKLFPKSD 300

```
Db 241 NRPIFGHYLDLHILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFOETARFQGVLEAP 359
Db 361 KELIFOETARFQGVLEAP 379

RESULT 7
US-10-701-490-8
; Sequence 8, Application US/10701490
; Publication No. US20040106141A1
; GENERAL INFORMATION:
; APPLICANT: PAUL S. MISCHER
; APPLICANT: CHARLES L. SAWYERS
; APPLICANT: BRADLEY L. SMITH
; APPLICANT: KATHERINE CROSBY
; TITLE OF INVENTION: METHODS AND MATERIALS FOR EXAMINING
; TITLE OF INVENTION: PATHWAYS ASSOCIATED WITH GLIOBLASTOMA PROGRESSION
; FILE REFERENCE: G&C 30435.148USU1
; CURRENT APPLICATION NUMBER: US/10/701,490
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/423,777
; PRIOR FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-701-490-8

Query Match 98.9%; Score 1859; DB 4; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPFEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPFEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Qy 121 VODLMTDLYKLLKSQOLSNNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
Db 121 VODLMTDLYKLLKSQOLSNNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Qy 241 NRPIFGHYLDLHILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLPKSD 300
Db 241 NRPIFGHYLDLHILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFOETARFQGVLEAP 359
Db 361 KELIFOETARFQGVLEAP 379

RESULT 8
US-10-735-118-3
; Sequence 3, Application US/10735118
; Publication No. US20040248151A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Bacus, Sarah S.
; APPLICANT: Smith, Bradley L.
; TITLE OF INVENTION: METHOD FOR PREDICTING THE RESPONSE TO HER2-DIRECTED THERAPY
; FILE REFERENCE: 6270-701-201
; CURRENT APPLICATION NUMBER: US/10/735,118
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 60/370,473
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/432,943
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 10/408,520
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-735-118-3

Query Match 98.9%; Score 1859; DB 5; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPFEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPFEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Qy 121 VODLMTDLYKLLKSQOLSNNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
Db 121 VODLMTDLYKLLKSQOLSNNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Qy 241 NRPIFGHYLDLHILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLPKSD 300
Db 241 NRPIFGHYLDLHILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFOETARFQGVLEAP 359
Db 361 KELIFOETARFQGVLEAP 379

RESULT 9
US-10-623-108-6
; Sequence 6, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-6

Query Match 98.9%; Score 1859; DB 5; Length 379;
```

Best Local Similarity 94.7%; Pred. No. 2.1e-145; Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGMWSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGMWSSAY 60

QY 61 DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

QY 241 NRPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAAWKLFPKSD 300
DB 241 NRPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAAWKLFPKSD 300

QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDFVABEPPTFAMELDDLPKRL 360

QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 10
US-10-945-684-2
; Sequence 2, Application US/10945684
; Publication No. US20050095657A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Methods and Kits for Detecting Proteins
; FILE REFERENCE: 50508-2280
; CURRENT APPLICATION NUMBER: US/10/945,684
; CURRENT FILING DATE: 2004-09-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-945-684-2

Query Match 98.9%; Score 1859; DB 5; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145; Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGMWSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGMWSSAY 60

QY 61 DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

QY 241 NRPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAAWKLFPKSD 300
DB 241 NRPIFPKGHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAAWKLFPKSD 300

QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDFVABEPPTFAMELDDLPKRL 360

QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 11
US-10-060-065-16
; Sequence 16, Application US/10060065
; Publication No. US20030017480A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/10/060,065
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-065-16

Query Match 98.6%; Score 1853; DB 4; Length 379;
Best Local Similarity 94.5%; Pred. No. 6.6e-145; Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGMWSSAY 60
DB 1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGPFDVGPRTQLOVIGEGAYGMWSSAY 60

QY 61 DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

Db 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240

Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVWAKLFPKSD 300

Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVWAKLFPKSD 300

Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKERL 340

Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 360

Qy 341 KELIFOETARFQPGVLEAP 359

Db 361 KELIFOETARFQPGVLEAP 379

RESULT 12

US-10-059-585-37

; Sequence 37, Application US/10059585

; Publication No. US20030082776A1

; GENERAL INFORMATION:

; APPLICANT: Ota, Toshio

; APPLICANT: Isogai, Takao

; APPLICANT: Nishikawa, Tetsuo

; APPLICANT: Hayashi, Koji

; APPLICANT: Otsuka, Kaoru

; APPLICANT: Yamamoto, Jun-ichi

; APPLICANT: Ishii, Shizuko

; APPLICANT: Sugiyama, Tomoyasu

; APPLICANT: Wakamatsu, Ai

; APPLICANT: Nagai, Keiichi

; APPLICANT: Otsuki, Tetsuji

; APPLICANT: Funahashi, Shin-Ichi

; APPLICANT: Senoo, Chiaki

; APPLICANT: Nezu, Jun-Ichi

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN

; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE

; FILE REFERENCE: 06501-098001

; CURRENT APPLICATION NUMBER: US/10/059,585

; CURRENT FILING DATE: 2002-01-29

; PRIOR APPLICATION NUMBER: PCT/JP00/05060

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: US 60/183,322

; PRIOR FILING DATE: 2000-02-17

; PRIOR APPLICATION NUMBER: US 60/159,590

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: JP 2000-118776

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: JP 2000-183767

; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 11-248036

; PRIOR FILING DATE: 1999-07-29

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 37

; LENGTH: 379

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-059-585-37

Query Match 98.6%; Score 1853; DB 4; Length 379;

Best Local Similarity 94.5%; Pred. No. 6.6e-145;

Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

Qy 1 MAAAAQGGGGGPRRTGEGVGPVGEVEMVKGPFDVGPRTQLOYIGEGAYGMVSSAY 60

Db 1 MAAAAQGGGGGPRRTGEGVGPVGEVEMVKGPFDVGPRTQLOYIGEGAYGMVSSAY 60

Qy 61 DHVKTRVAKKISPEHQTQCORTREIQILLRFHENVIGIRDLIRASTLEAMRDVVI 120

Db 61 DHVKTRVAKKISPEHQTQCORTREIQILLRFHENVIGIRDLIRASTLEAMRDVVI 120

Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLNTTCDL 180

Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLNTTCDL 180

Qy 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240

Db 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240

Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVWAKLFPKSD 300

Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVWAKLFPKSD 300

Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKERL 340

Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 360

Qy 341 KELIFOETARFQPGVLEAP 359

Db 361 KELIFOETARFQPGVLEAP 379

RESULT 13

US-10-170-663-25

; Sequence 25, Application US/10170663

; Publication No. US20030165899A1

; GENERAL INFORMATION:

; APPLICANT: Su, Michael Shin-San

; Fox, Ted

; Wilson, Keith Phillip

; Germann, Ursula A.

; TITLE OF INVENTION: Methods For Designing Inhibitors of

; Serine/Threonine Kinases and Tyrosine Kinase

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Neave

; STREET: 1251 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10020

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/170,663

; FILING DATE: 12-Jun-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/025,580

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Haley, James F.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: VPI 97-104

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 596-9000

; TELEFAX: (212) 596-9090

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 379 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-10-170-663-25

Query Match 98.6%; Score 1853; DB 4; Length 379;

Best Local Similarity 94.5%; Pred. No. 6.6e-145;

Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

Qy 1 MAAAAQGGGGGPRRTGEGVGPVGEVEMVKGPFDVGPRTQLOYIGEGAYGMVSSAY 60

Db 1 MAAAAAQQGGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Qy 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDINSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDINSVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLPPKSD 300
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLPPKSD 300
Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPFFAMELDDLPKRL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVABEPTFFAMELDDLPKRL 360
Qy 341 KELIFOETARFQPGVLEAP 359
Db 361 KELIFOETARFQPGVLEAP 379

RESULT 14
US-10-893-072-25
; Sequence 25, Application US/10893072
; Publication No. US20040259166A1
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; Fox, Ted
; Wilson, Keith Phillip
; Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/893,072
; FILING DATE: 15-Jul-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-893-072-25
Query Match 98.6%; Score 1853; DB 5; Length 379;
Best Local Similarity 94.5%; Pred. No. 6.e-145; Indels 20; Gaps 1;
Matches 358; Conservative 0; Mismatches 1;
Qy 1 MAAAAAQQGGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Qy 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDINSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDINSVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLPPKSD 300
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLPPKSD 300
Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPFFAMELDDLPKRL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVABEPTFFAMELDDLPKRL 360
Qy 341 KELIFOETARFQPGVLEAP 359
Db 361 KELIFOETARFQPGVLEAP 379

RESULT 15
US-10-072-036-39
; Sequence 39, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJRON
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 39
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-10-072-036-39

Query Match 98.6%; Score 1853; DB 4; Length 631;
Best Local Similarity 94.5%; Pred. No. 1.2e-144;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;
Qy 1 MAAAAAQQGGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Db 253 MAAAAAQQGGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 312
Qy 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI 120

313	Db	DHVRKTRVAIKKISPEFHQTYCQRTREIQILLRFRHENVIGIRDLIRASTLEARNRDVYI	372
121	QY	VQDLMETDLYKLLKSKQQLSNDHIYCVFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL	180
373	Db	VQDLMETDLYKLLKSKQQLSNDHIYCVFLYQILRGLKYIHSANVLHRDLKPSNLLSNTTCDL	432
181	QY	KICDFGLARIADPEHDHTGFTTEYVATRWYRAPEIMLNSKGYTKSIDIVSWGCIILAEMLS	240
433	Db	KICDFGLARIADPEHDHTGFTTEYVATRWYRAPEIMLNSKGYTKSIDIVSWGCIILAEMLS	492
241	QY	NRPIPPGKHYLDQLNHIILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAKLFPKSD	300
493	Db	NRPIPPGKHYLDQLNHIILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAKLFPKSD	552
301	QY	SKAULDLLDRMLTTFNENKRIIT-----VAEEFFTFAMELDDDLPKERL	340
553	Db	SKAULDLLDRMLTTFNENKRIITVTEALAHPPYLEQYDPTDEPVAEEPTFAMELDDDLPKERL	612
341	QY	KELIFQETARFQPGVLEAP	359
613	Db	KELIFQETARFQPGVLEAP	631

Search completed: February 6, 2006, 15:40:57
Job time : 126.184 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:37:06 ; Search time 10.8631 Seconds
(without alignments)
387.254 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAGGGGGPRRTGV.....LKELIFQTARFGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA_New:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1853	98.6	379	7	US-11-109-156-16
2	1538.5	81.9	360	6	US-10-878-556A-134
3	1538.5	81.9	360	7	US-11-186-284-115
4	758.5	40.4	360	6	US-10-886-329-1
5	704.5	37.5	365	6	US-10-770-726-69
6	629	33.5	422	6	US-10-857-780-21
7	629	33.5	422	7	US-11-127-817-18
8	629	33.5	426	7	US-11-127-817-20
9	629	33.5	464	7	US-11-127-817-19
10	491	26.1	460	7	US-11-024-959-384
11	490.5	26.1	294	7	US-11-024-959-263
12	481.5	25.6	294	7	US-11-024-959-265
13	479.5	25.5	304	7	US-11-024-959-264
14	477.5	25.4	302	7	US-11-024-959-262
15	472.5	25.1	520	7	US-11-024-959-272
16	471.5	25.1	298	6	US-10-770-726-51
17	458	24.4	277	7	US-11-127-817-21
18	451.5	24.0	346	6	US-10-770-726-55
19	450.5	24.0	297	7	US-10-770-726-48
20	450.5	24.0	297	7	US-11-109-156-11
21	449.5	23.9	795	6	US-10-770-726-49
22	449	23.9	496	6	US-10-770-726-72
23	448.5	23.9	555	7	US-11-024-959-270
24	445.5	23.7	509	7	US-11-024-959-393
25	443.5	23.6	463	7	US-11-024-959-395

26	442	23.5	292	6	US-10-770-726-53	Sequence 53, Appl
27	441	23.5	483	6	US-10-451-375-12	Sequence 12, Appl
28	439	23.4	476	7	US-11-024-959-385	Sequence 385, App
29	438	23.3	433	6	US-10-770-726-65	Sequence 65, Appl
30	437	23.3	313	7	US-11-024-959-267	Sequence 267, App
31	433	23.0	706	7	US-11-024-959-269	Sequence 269, App
32	431	22.9	477	7	US-11-024-959-261	Sequence 261, App
33	426.5	22.7	330	6	US-10-786-065-8	Sequence 8, Appl
34	424.5	22.6	330	6	US-10-786-065-9	Sequence 9, Appl
35	423.5	22.5	330	6	US-10-786-065-10	Sequence 10, Appl
36	421	22.4	583	7	US-11-024-959-268	Sequence 268, App
37	419	22.3	435	6	US-10-786-065-5	Sequence 5, Appl
38	419	22.3	718	7	US-11-024-959-273	Sequence 273, App
39	415	22.1	302	7	US-11-024-959-386	Sequence 386, App
40	414	22.0	845	7	US-11-024-959-391	Sequence 391, App
41	412	21.9	469	7	US-11-024-959-271	Sequence 271, App
42	405	21.6	797	7	US-11-024-959-392	Sequence 392, App
43	401.5	21.4	303	6	US-10-770-726-52	Sequence 52, Appl
44	398.5	21.2	305	7	US-11-024-959-266	Sequence 266, App
45	395	21.0	302	7	US-11-024-959-387	Sequence 387, App

ALIGNMENTS

RESULT 1
US-11-109-156-16
; Sequence 16, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isegai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT FILING DATE: 2005-04-19
; PRIOR FILING DATE: 2002-01-29
; PRIOR FILING DATE: 2002-01-29
; PRIOR FILING DATE: 2000-07-28
; PRIOR FILING DATE: 1999-10-18
; PRIOR FILING DATE: 2000-02-17
; PRIOR FILING DATE: 2000-02-17
; PRIOR FILING DATE: 1999-07-29
; PRIOR FILING DATE: 2000-01-11
; PRIOR FILING DATE: 2000-01-11
; PRIOR FILING DATE: 2000-05-02
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-16

Query Match		98.6%;	Score 1853;	DB 7;	Length 379;
Best Local Similarity		94.5%;	Pred. No. 4.4e-157;		
Matches		358;	Conservative	0;	Mismatches 1; Indels 20; Gaps 1;
QY	1	MAAAAQGGGGPRTEGPGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAQGGGGPRTEGPGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
QY	61	DHVKTRVAIKKISPEHOTYQORTLEIRIQILLRFRHENVIGIRDILRASTLEAMRDVYI	120		
DB	61	DHVKTRVAIKKISPEHOTYQORTLEIRIQILLRFRHENVIGIRDILRASTLEAMRDVYI	120		
QY	121	VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	180		
DB	121	VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	180		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	300		
DB	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	300		
QY	301	SKALDILLDRMLTFNPNKRIT-----VAEPTTFAMELDDLPKERL	340		
DB	301	SKALDILLDRMLTFNPNKRITVEBALAHPLYEQYDPTDFVAEPTTFAMELDDLPKERL	360		
QY	341	KELIFOETARFQGVLEAP	359		
DB	361	KELIFOETARFQGVLEAP	379		
RESULT 2					
US-10-878-556A-134					
; Sequence 134, Application US/10878556A					
; Publication No. US20050266399A1					
; GENERAL INFORMATION:					
; APPLICANT: Hoffmann La-Roche Inc.					
; TITLE OF INVENTION: HCV regulated protein expression					
; FILE REFERENCE: 21762					
; CURRENT APPLICATION NUMBER: US/10/878,556A					
; CURRENT FILING DATE: 2004-06-28					
; NUMBER OF SEQ ID NOS: 199					
; SOFTWARE: PatentIn version 3.1					
; SEQ ID NO 134					
; LENGTH: 360					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
; PUBLICATION INFORMATION:					
; DATABASE ACCESSION NUMBER: sw_hum/mk01_human					
; DATABASE ENTRY DATE: 1992-12-01					
US-10-878-556A-134					
Query Match		81.9%;	Score 1538.5;	DB 6;	Length 360;
Best Local Similarity		79.1%;	Pred. No. 3.3e-129;		
Matches		296;	Conservative	18;	Mismatches 23; Indels 37; Gaps 2;
QY	1	MAAAAQGGGGPRTEGPGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAAGAG-----PEMVRQGVDPGPRYTNSYIGEGAYGMVCSAY	43		
QY	61	DHVKTRVAIKKISPEHOTYQORTLEIRIQILLRFRHENVIGIRDILRASTLEAMRDVYI	120		
DB	44	DNVKNRVAIKKISPEHOTYQORTLEIRIKILLRFRHENIIGINDIRAPTIEQMKDVI	103		
QY	121	VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	180		
DB	104	VQDLMETDLYKLLKTOHLSNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	163		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	164	KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	223		

QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	300		
DB	224	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	283		
QY	301	SKALDILLDRMLTFNPNKRITV-----AEPPTTFAMELDDLPKERL	340		
DB	284	SKALDILLDRMLTFNPNKRITVEQALAHPLYEQYDPSDEPTAEAPFFEDMELDDLPKERL	343		
QY	341	KELIFOETARFQGP	354		
DB	344	KELIFEETARFQGP	357		
RESULT 3					
US-11-186-284-115					
; Sequence 115, Application US/11186284					
; Publication No. US20050266493A1					
; GENERAL INFORMATION:					
; APPLICANT: Millennium Pharmaceuticals, Inc.					
; APPLICANT: Berger, Allison					
; APPLICANT: Guillemette, Tracy L.					
; APPLICANT: Kamatkar, Shubhangi					
; APPLICANT: Schlegel, Robert					
; APPLICANT: Monahan, John E.					
; APPLICANT: Thibodeau, Stephen N.					
; APPLICANT: Burgart, Lawrence J.					
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND					
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND					
; TITLE OF INVENTION: THERAPY OF COLON CANCER					
; FILE REFERENCE: MP001-029P2RNM					
; CURRENT APPLICATION NUMBER: US/11/186,284					
; CURRENT FILING DATE: 2005-07-21					
; PRIOR APPLICATION NUMBER: US/10/301,822					
; PRIOR FILING DATE: 2002-11-21					
; PRIOR APPLICATION NUMBER: US 60/339,971					
; PRIOR FILING DATE: 2001-12-10					
; PRIOR APPLICATION NUMBER: US 60/361,978					
; PRIOR FILING DATE: 2002-03-05					
; PRIOR APPLICATION NUMBER: US 60/381,988					
; PRIOR FILING DATE: 2002-05-20					
; NUMBER OF SEQ ID NOS: 228					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 115					
; LENGTH: 360					
; TYPE: PRT					
; ORGANISM: Homo Sapiens					
US-11-186-284-115					
Query Match		81.9%;	Score 1538.5;	DB 7;	Length 360;
Best Local Similarity		79.1%;	Pred. No. 3.3e-129;		
Matches		296;	Conservative	18;	Mismatches 23; Indels 37; Gaps 2;
QY	1	MAAAAQGGGGPRRTGPGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAAGAG-----PEMVRQGVDPGPRYTNSYIGEGAYGMVCSAY	43		
QY	61	DHVKTRVAIKKISPEHOTYQORTLEIRIQILLRFRHENVIGIRDILRASTLEAMRDVYI	120		
DB	44	DNVKNRVAIKKISPEHOTYQORTLEIRIKILLRFRHENIIGINDIRAPTIEQMKDVI	103		
QY	121	VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	180		
DB	104	VQDLMETDLYKLLKTOHLSNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL	163		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	164	KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	223		
QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	300		
DB	224	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	283		

QY	241	NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	300
DB	224	NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	283
QY	301	SKALDLDLDRMLTNPNNKRTIV-----AEPFTFAMELDDLPKERL	340
DB	284	SKALDLDLDRMLTNPNNKRIEVEQALAHPLYEQYYDSDFEPIAEAPPFDMELDDLPKEKL	343
QY	341	KELIFQETARFQPG 354	
DB	344	KELIFQETARFQPG 357	
RESULT 3			
US-11-186-284-115			
; Sequence 115, Application US/11186284			
; Publication No. US20050266493A1			
; GENERAL INFORMATION:			
; APPLICANT: Millennium Pharmaceuticals, Inc.			
; APPLICANT: Berger, Allison			
; APPLICANT: Guillemette, Tracy L.			
; APPLICANT: Kamatkar, Shubhangi			
; APPLICANT: Schlegel, Robert			
; APPLICANT: Monahan, John E.			
; APPLICANT: Thibodeau, Stephen N.			
; APPLICANT: Burgart, Lawrence J.			
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND			
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
; FILE REFERENCE: MPM01-029P2RNM			
; CURRENT APPLICATION NUMBER: US/11/186,284			
; CURRENT FILING DATE: 2005-07-21			
; PRIOR APPLICATION NUMBER: US/10/301,822			
; PRIOR FILING DATE: 2002-11-21			
; PRIOR APPLICATION NUMBER: US 60/339,971			
; PRIOR FILING DATE: 2001-12-10			
; PRIOR APPLICATION NUMBER: US 60/361,978			
; PRIOR FILING DATE: 2002-03-05			
; PRIOR APPLICATION NUMBER: US 60/381,988			
; PRIOR FILING DATE: 2002-05-20			
; NUMBER OF SEQ ID NOS: 228			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 115			
; LENGTH: 360			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-11-186-284-115			
Query Match 81.9%; Score 1538.5; DB 7; Length 360;			
Best Local Similarity 79.1%; Pred. No. 3.3e-129;			
Matches 296; Conservative 18; Mismatches 23; Indels 37; Gaps 2;			
QY	1	MAAAAQGGGGPRTEGVGPGEVEMVKQPDVGBPRYTQLOYIGEGAYGMVSSAY	60
DB	1	MAAAAAGAG-----PEMVRQGVDPGPRYTNSYIGEGAYGMVCSAY	43
QY	61	DHVKTRVAIKKISPEHOTYQORTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI	120
DB	44	DNVKNRVAIKKISPEHOTYQORTREIKILLRFRHENIIGINDIRAPTIEQMVDVYI	103
QY	121	VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRLGKVIHSAANVLRDLKPSNLLINTTCDL	180
DB	104	VQDLMETDLYKLLKTOHLSNDHCYFLYQILRLGKVIHSAANVLRDLKPSNLLINTTCDL	163
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240
DB	164	KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	223
QY	241	NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	300
DB	224	NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD	283

```
QY 301 SKALDLDRLMTNPNKRTV-----AEPFTFAMELDDLKPERL 340
Db 284 SKALDLDKMLTNPBKRIEVEQALAHVLEQYDFSDPEIAEAPKPFDMELDLKPERL 343
QY 341 KELIFORTARFQPG 354
Db 344 KELIFETARFQPG 357

RESULT 4
US-10-886-329-1
; Sequence 1, Application US/10886329
; Publication No. US20050288286A1
; GENERAL INFORMATION:
; APPLICANT: Deciphra Pharmaceuticals, Inc.
; APPLICANT: Flynn, Daniel L
; APPLICANT: Pettillo, Peter A
; TITLE OF INVENTION: Anti-Inflammatory Medicaments
; FILE REFERENCE: 34477-CIP
; CURRENT APPLICATION NUMBER: US/10/886,329
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US 60/437,415
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US 10/746,460
; PRIOR FILING DATE: 2003-12-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-886-329-1

Query Match 40.4%; Score 758.5; DB 6; Length 360;
Best Local Similarity 45.5%; Pred. No. 5.2e-60;
Matches 156; Conservative 63; Mismatches 97; Indels 27; Gaps 4;

QY 36 FDVGPRTQIYIGEGAGYGVSSAYDHRKTRVAIKKIS-PPEHOTYQORTLREIQILLR 94
Db 18 WEVPERYQNLSPVSGAYGVCVCAAFDTKGLRVAVKLSRPFQSIHAKRTYRELRLKH 77
QY 95 FPHENVIGIRDLR-ASTLEAMRDVYVQDLMETDLYKLLKSQOLNSDHCYFLYQILRG 153
Db 78 MKHENVIGLDDVTPARSLDEFNFVYVTLVLMGADLNNIVKQCLTDDHVFQILYQILRG 137
QY 154 LKVIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAP 213
Db 138 LKVIHSADIIRHLKPSNLAVNDCDELKILDFGLARHTDE-----MTGYVATRWYRAP 191
QY 214 EIMLNSKGYTKSIDIWSVGCILAEMLSNRPIFFKGHYLDQNLHILGILGSPSQEDLNCII 273
Db 192 EIMLNMHYNQTVDIWSVGCIMAEELLTGRTLFFGTDHIDQLKILRLVGTGPAELKKIS 251
QY 274 NMKARNYQSLPSKTKVAVAKLPKSDSKALDLDRLMTNPNKRTVAB-----323
Db 252 SESARNYQSLTQMPKMFANVFIGANPLAVDLLEKMLVLDSDKRLTAAQALAHAYFAQY 311
QY 324 -----EPFTFAMELDDLKPERLKELIFQETARFQPGVLE 357
Db 312 HDPDDEPVADPYQDSFESRDLLEKMLVLDSDKRLTAAQALAHAYFAQY 354

RESULT 5
US-10-770-726-69
; Sequence 69, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)
```

```
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 69
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-69

Query Match 37.5%; Score 704.5; DB 6; Length 365;
Best Local Similarity 42.8%; Pred. No. 3.3e-55;
Matches 143; Conservative 64; Mismatches 98; Indels 29; Gaps 5;

QY 42 YTQLOVIGEGAGYGVSSAYDHRKTRVAIKKIS-PPEHOTYQORTLREIQILLRPHENV 100
Db 25 YVSPTHVSGAYGVCVCAIDKRSGEKVAIKLSRPFQSEIFAKRAYRELLLKHMQHENV 84
QY 101 IGRIDILR-ASTLEAMRDVYVQDLMETDLYKLLKSQOLNSDHCYFLYQILRGKLYIHS 159
Db 85 IGLLDVFTPASSLRNFYDFLYMPFMQTDLQIM-GMEFSEBKIQILVYQMLKGLKLYIHS 143
QY 160 ANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLS 219
Db 144 AGVHRDLKPSNLAVNDCDELKILDFGLARHADAE-----MTGYVATRWYRAPEVILSW 197
QY 220 KGYTKSIDIWSVGCILAEMLSNRPIFFKGHYLDQNLHILGILGSPSQEDLNCIINKARN 279
Db 198 MHNQTVDIWSVGCIMAEMLTGTLPFGKDYLDQLTQILKVTGPGTEFVKLNDKAAS 257
QY 280 YLQSLPSKTKVAVAKLPKSDSKALDLDRLMTNPNKRTV-----321
Db 258 YIQSLPQTPRKDFTQLFPRASPOAADLLEKMLELDVDRKRLTAAQALAHAYFAQY 317
QY 322 --AEPFTFAMELDDLKPERLKELIFQETARFOP 353
Db 318 TEAQPFDSLEHEKLTVDENKQHYKEIVNFSF 351
```

```
RESULT 6
US-10-857-780-21
; Sequence 21, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE REFERENCE: THEREOF
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-780-21

Query Match 33.5%; Score 629; DB 6; Length 422;
Best Local Similarity 39.7%; Pred. No. 2e-48;
Matches 138; Conservative 72; Mismatches 94; Indels 44; Gaps 10;
```


Job time : 11.8631 secs

```
Db 63 IVALYDVWHTTE-----KGLTLVFEFLDQDLKYLDA CGNGLEPYTVKSFYQLQGIAP 117
Qy 157 IHSANVLHRLDKPSNLLINTTCDLKICDFGLAR-IADPEHDHGTGFLTEYVATWYRAPEI 215
Db 118 CHEHRLVHRLDKFQNLINWEGSLKGLADFGARAFGIPVRYN-----THEVTLWYRAPDV 173
Qy 216 MLNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYDQLNHLGILGSPSQEDLNCIINN 275
Db 174 LMGRKYSTQVDIWSVGCIFAEMVNGRFLPPGSSEODQLLRIFKTLGTPLSKTWPMAEL 233
Qy 276 KARYLQSLPSKYKVAWAKLFPKS-DSKALDLDRLMLTFNPNGRITVAE 323
Db 234 P--DFKDNFPKYVQVSPKKICPKKLDKTKGLDLSRMLQYDPAKRISABQ 280

RESULT 15
US-11-024-959-272
; Sequence 272, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; PRIOR FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 272
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-272

Query Match 25.1%; Score 472.5; DB 7; Length 520;
Best Local Similarity 34.2%; Pred. No. 2e-34;
Matches 122; Conservative 56; Mismatches 116; Indels 63; Gaps 8;

Qy 42 YTOIQYIGEGAYGVSSAYDHRKTRVAIKKI-SPPEHQTYCQRTLRIOILLRFRHENV 100
Db 25 FEKLEQIGEGTYGOVYMAKEKKTGEI VALKKIRMDNEREGFPITAI REIKLKKLHHENV 84
Qy 101 IGIRDILRASTLEAMRD-----VTVQDLMETDLYKLL--KSQQLSNDHICYFL 147
Db 85 IKLKEIVTSPGPEKDEQREGPNKYGGIYVFFYMDHDLTGLADRFGRFSPQIKCYM 144
Qy 148 YQILRGLKYIHSANVLHRLDKPSNLLINTTCDLKICDFGLARIADPEHDHGTGFLTEYVAT 207
Db 145 RQLLTGLHYCHINQVLRHDIKGNSLLIDNEGNLKLDADFGLAR--SFSNDHNANITNRVIT 202
Qy 208 RWTAPRIMLNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYDQLNHLGILGSPSQE 267
Db 203 LWYRPPPELLLGATKYGPAVDMWSVGCIFAELLHGKPIPPGKDEPEQLNKIFELCGAPDE- 261
Qy 268 DLNCIINWKARNYLQSLPSKYKVAW-----AKLFPKSDSKALDLDRLMLTF 313
Db 262 -----INW-----PGVSKI PWYNNFKTRPMKRRLREVFHRHLELLERMLTL 307
Qy 314 NPKRITVAEBPFTFAMELDDLP-----KERLKELIPOETARFQ 352
Db 308 DPSORISAKDALDAEYFWADPLPCDPKSLPKYESSHEFPQTKKKGQQORQHEETAKRQ 364
```

Search completed: February 6, 2006, 15:41:23

This Page Blank (uspto)